

Sequence	Strd	zScore	EScore	len	Documentation	
P1r2:DB1986	+	1580.00	2425.73	1.7e-127	298	probable acyltransferase NMA06666
P1r2:DB11040	+	1577.00	2421.11	3.1e-127	298	HtrB/MSDB family protein NMB18108
P1r2:CG92691	+	263.50	395.83	1.9e-14	316	Lipid A biosynthesitis lauroyl ac
P1r2:HB3662	+	253.00	389.50	4.5e-14	295	probable 2-OH-lauroyltransferas
P1r2:DB3239	+	228.50	311.99	1.9e-11	312	probable lauroyl acyltransferas
P1r2:SI16888	+	215.00	321.35	2.7e-10	306	Lipid A biosynthesitis lauroyl ac
P1r2:DB5667	+	215.00	321.35	2.7e-10	306	Lipid A biosynthesitis lauroyl ac
P1r2:HB0807	+	215.00	321.35	2.7e-10	306	Lipid A biosynthesitis lauroyl ac
P1r2:AT0841	+	214.00	319.81	3.3e-10	306	Lipid A biosynthesitis lauroyl ac
P1r2:DB1005	+	204.50	305.67	2.1e-09	289	probable membrane protein YP038
P1r2:DB1857	+	203.50	304.13	2.6e-09	289	HtrB/MSDB family protein NMB1441
P1r2:AE0807	+	199.50	297.46	5.8e-08	306	probable acetyltransferase NMA1
P1r2:EB2847	+	196.50	293.65	1.0e-08	279	probable acyltransferase (EC 2
P1r2:BP1036	+	189.50	281.43	4.2e-08	328	Lipid A biosynthesitis lauroyl ac
P1r2:G65011	+	189.50	281.43	4.2e-08	328	probable heat shock protein (im
P1r2:DB5880	+	189.50	281.43	4.2e-08	328	probable lauroyl acyltransferas
P1r2:AD0637	+	185.00	275.11	1.0e-07	306	probable heat shock protein ddd
P1r2:164053	+	178.50	264.75	3.7e-07	318	Lipid A biosynthesitis lauroyl ac
P1r2:ET1569	+	177.50	252.34	1.3e-06	455	membrane-bound lytic transglyc
P1r2:ET1631	+	169.50	251.69	1.3e-06	455	probable acyltransferase - ChlA
P1r2:PB1730	+	166.50	240.05	6.2e-06	290	Lipid A biosynthesitis lauroyl ac
P1r2:A60251	+	156.50	233.86	1.9e-05	453	Lipid A biosynthesitis lauroyl ac
P1r2:AA3391	+	152.50	224.98	6.3e-05	320	Lipid A biosynthesitis acyltransf
P1r2:T42129	+	151.00	220.09	8.6e-05	343	Lipid A biosynthesitis lauroyl ac
P1r2:CG7836	+	149.00	220.09	0.0001	290	probable acyltransferase (EC 2
P1r2:DB4137	+	148.50	218.59	0.0001	315	hypothetical protein htrB (impo
P1r2:EB5134	+	147.50	212.48	0.0002	315	Lipid A biosynthesitis lauroyl ac
P1r2:EB1551	+	146.00	211.36	0.0002	529	hypothetical protein [imported]
P1r2:DB72119	+	146.00	211.26	0.0002	462	Lipid A biosynthesitis lauroyl ac
P1r2:D97552	+	141.00	207.11	0.0006	467	acyltransferase - Chlamydomo
P1r2:AF2772	+	141.00	207.11	0.0006	312	Lipid A biosynthesitis lauroyl ac
P1r2:EB2350	+	141.00	206.75	0.0006	312	Lipid A biosynthesitis lauroyl ac
P1r2:G82350	+	136.50	200.01	0.0015	325	Lipid A biosynthesitis lauroyl ac
P1r2:Y43023	+	133.50	194.15	0.0028	318	Lipid A biosynthesitis (Kdo)2-(4
P1r2:T057722	+	133.00	193.15	0.0030	405	Lipid A biosynthesitis lauroyl ac
P1r2:EB90949	+	129.50	189.08	0.0060	335	cecoA protein - Burkholderia ceg
P1r2:BA85798	+	129.50	189.08	0.0060	323	lipoxygenase (EC 3.2.1.73) - bar
P1r2:J000405	+	129.00	177.46	0.0078	1106	heat shock protein MsbA [import
P1r2:BA42668	+	127.50	186.00	0.0089	323	suppressor of htrB, heat shock
P1r2:ME0742	+	127.50	186.00	0.0089	323	hypothetical I19.5K protein (u
	+					(Kdo)2-(lauroyl)-Lipid IV4 ac
	+					Lipid A acyltransferase (EC 2.3

p1r2:C52123	+ 123.50	175.58	0.0160	860	! probable endoglucanase - Rum
p1r2:C507132	+ 125.00	170.93	0.0223	866	! hypothetical protein 1 - Rh
p1r2:C75580	+ 116.00	166.93	0.0880	376	! adenine deaminase-related p
p1r2:C68695	+ 115.50	167.58	0.0951	330	! conserved hypothetical prote
p1r2:C93193	+ 115.00	164.06	0.1094	437	! hypochlorite protein 3 (gylr
p1r2:A45093	+ 114.50	160.45	0.1259	603	! hypochlorite protein [impor
p1r2:A87528	+ 114.00	165.43	0.1277	314	! conserved hypothetical prot
p1r2:T50711	+ 114.00	160.19	0.1380	569	! ucase (EC 3.5.1.5) alpha ch
p1r2:A34615	+ 114.00	159.37	0.1397	75	! profilaggrin - rat (fragmen
p1r2:E87627	+ 113.00	156.36	0.1739	628	! hypochlorite protein CC3055
p1r2:T42761	+ 113.00	151.60	0.2165	338	! Bassoon protein - rat
p1r2:I13869	+ 112.00	141.27	0.2234	1104	! transcription factor NFκB1
p1r2:T01480	+ 112.00	151.12	0.2239	1122	! exit protein - Mycobacteri
p1r2:T02345	+ 112.00	147.00	0.2380	1791	! hypochlorite protein KIAA0
p1r2:A35638	+ 111.50	144.23	0.2706	2248	! profilaggrin - human (frag
p1r2:H71934	+ 111.00	140.42	0.3232	328	! probable lipid A biosynthes
p1r2:B97369	+ 109.00	148.32	0.3944	913	! translation initiation facto
p1r2:A82587	+ 109.00	147.03	0.3944	1057	! ribonuclease E PA2976 [imp
p1r2:H83373	+ 108.00	145.99	0.4862	988	! probable two-component sens
p1r2:D84595	+ 107.50	147.76	0.5169	748	! PEARR1 4 protein [importe
p1r2:C83372	+ 107.00	154.26	0.5124	328	! heat shock protein B - Heli
p1r2:I150620	+ 106.50	153.03	0.5218	377	! hypochlorite protein PA2181
p1r2:T47370	+ 105.00	142.25	0.6680	1173	! prokri2 - chicken (fragme
p1r2:A81318	+ 104.50	151.26	1.05	3342	! Bassoon protein - mouse
p1r2:JH0501	+ 104.50	146.53	0.8284	295	! probable lipid A biosynthes
p1r2:T030251	+ 104.50	139.50	0.8895	509	! zinc finger protein CKR1 - c
p1r2:S536149	+ 104.50	137.00	0.9872	1130	! repeat - mouse
p1r2:T24818	+ 104.50	134.37	1.02	1500	! copper-transferring AtPase
p1r2:S27953	+ 104.00	151.45	1.07	2022	! glucan 1,4-alpha-glucosida
p1r2:H87520	+ 104.00	147.37	0.9026	267	! Mx protein - pig (fragment)
p1r2:T31421	+ 103.50	137.63	0.9588	424	! hypochlorite protein CC2193
p1r2:B87571	+ 103.50	147.22	1.14	1173	! C-terminal domain-binding f
p1r2:AEL1000	+ 103.00	145.04	1.21	362	! hypochlorite protein CC2597
p1r2:A55360	+ 102.50	146.02	1.29	425	! Damx protein [imported] - S
p1r2:A61183	+ 102.00	140.02	1.32	688	! zinc finger protein (clone J
p1r2:IEDE1F	+ 101.50	146.22	1.51	312	! hypochlorite protein (psdb i
p1r2:E97631	+ 101.50	132.62	1.85	1460	! immediate-early protein IEI
p1r2:T07339	+ 101.00	141.73	1.76	476	! probable membrane transport
p1r2:A69046	+ 100.50	144.90	1.86	341	! plastocyanin-plastocyanin
p1r2:T30136	+ 100.50	139.15	2.00	584	! conserved hypochlorite prot
p1r2:UCS280	+ 100.50	127.44	2.03	654	! hypochlorite protein ClC411
p1r2:T18312	+ 99.50	130.70	2.28	2206	! voltage-dependent calcium
p1r2:T13422	+ 99.50	129.46	2.75	1279	! hypochlorite protein L7610
p1r2:T33997	+ 99.00	138.74	2.61	471	! C-terminal domain-binding f
p1r2:B75310	+ 99.00	137.73	2.65	528	! hypochlorite protein W0361
p1r2:SA4152	+ 98.00	142.55	2.90	229	! conserved hypochlorite prot
p1r2:CG71529	+ 98.00	127.62	3.67	1396	! Microneme antigen - Sarcocy
p1r2:T35574	+ 97.50	138.07	3.43	391	! DNA-directed RNA polymeras
p1r2:JH0163	+ 97.50	132.96	3.70	698	! probable alanine racemase -
p1r2:JH0162	+ 97.50	133.93	3.70	700	! No-on-transient A protein fo
p1r2:JC4161	+ 97.50	131.48	3.78	825	! No-on-transient A protein fo
p1r2:T00062	+ 97.50	125.81	4.11	1571	! DNA-binding protein 5E5 - r
p1r2:JG4743	+ 97.50	120.72	4.43	2796	! hypochlorite protein KIAA0
p1r2:H84824	+ 97.00	130.82	3.92	510	! fatty-acid synthase (EC 2.3.
p1r2:B30843	+ 97.00	134.95	4.17	815	! En/Spm-like transposon prote
p1r2:JH0689	+ 97.00	130.8			

A:Accession: D81986  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-298 <PKR>  
 A:Cross-references: GB:AL162753; GB:AL157959; MID:g7379120; PIDN:CAB83949.1; PID:g737939  
 A:Experimental source: serogroup A, strain Z2491  
 C:Genetics:  
 A:Gene: NMA0662

alignment\_scores:  
 Quality: 1580.00 Length: 298  
 Ratio: 5.302 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-303-518D-569 x D81986 ..

Align seg 1/1 to: D81986 from: 1 to: 298

```

1 ATGTTTCGTTTACATTCAGCGCTTTCCCTTGGCAACCCCATGCA 50
|||||
1 MetPheArgLeuGlnPheArgLeuPheProPheLeuArgThrAlaMetHis 17
51 CATCCGTGTGACCGCCCTGCTCAAAATGCTCTCCCTGCGCGCTTCT 100
|||||
17 sileLeuLeuThrAlaLeuLeuLysCysLeuSerLeuLeuProLeuSerc 34
101 GTCTGCACACGCTGGGAAACCGGCTCGACATCTGGCGTTTACCTTTA 150
|||||
34 yLeuHisThrLeuGlnLysnArgLeuGlnHisLeuAlaPheThrLeuLeu 50
151 AAGGAAGACCGCGCGCATGCTGCCCAATATGCTCAGCGAGCATGAA 200
|||||
51 LysGlnAspArgAlaArgIleValAlaAsnMetArgGlnAlaGlyMetAs 67
67 nProAspProLysThrValLysAlaValPheAlaGlnThrAlaLysGly 84
251 GTTTGGAACTTCCCCCGCTTTTTCAGAAAACCGGAAGACATAGAAACA 300
|||||
84 LysGlnLeuAlaPheAlaPheArgLysProGlnAspIleGlnThr 100
301 AAGTTCAAGCGGTACAGCGCTGGGAACTGTGCAGACGCTTGCACAA 350
|||||
101 MetPheLysAlaValHisGlyTrpIleHisValGlnGlnAlaLeuAspLys 117
351 ACAGCAAGGGCTGTATTCATCAGCCCGCACATCGGCAAGCTACGATTGG 400
|||||
117 sHisGlnGlyLeuLeuPheIleThrProHisIleGlySerTyrAspLeuG 134
401 GGGGACGCTACATCAGCCGACGAGCTTCCTGCCGTGACCGCCATGAC 450
|||||
134 LysIleArgGlyTrpIleSerGlnGlnLeuProPheProLeuThrAlaMetYtr 150
451 AAACCGCCGAAATCAAGCGCATAGCAAAATCATGACAGCGGCGACGAT 500
|||||
151 LysProProLysIleLysAlaIleAspLysIleMetGlnAlaGlyArgVal 167
501 TCGCGGCAAGGAAACCGCGCTACACGATACAGAGGGGTCAACAA 550
|||||
167 LArgGlyLysGlyLysThrAlaProThrSerIleGlnGlyValLysGlnI 184
551 TCATCAAGACCGCTGGTGGGCGGAAAGCAACATCGTCTCTGCCGACAC 600
|||||
184 IeLeuLysAlaLeuArgSerGlyGlnAlaThrIleValLeuProAspHis 200
601 GTCCCTCCCTCAAGAAAGCGGCGGAAAGCGATGAGGTGATTTCTTCG 650
|||||
201 ValProSerProGlnGlnGlyGlyGlnGlyValITrpValAspPhePhe 217
651 CAACCTGCTTATACCATGACGCTGGCGGCAAAATTTGGCACACGTCAAAG 700

```

seq\_name: plr2:B81040

seq\_documentation\_block:

HtB/MSB family protein NMB1801 [imported] - Neisseria meningitidis (strain MC58 ser C);Species: Neisseria meningitidis  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C:Accession: B81040  
 R:Tetzelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, R.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.; Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizzza, M. Science 287, 1809-1815, 2000  
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R. A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. A:Reference number: A81000; MUID:20175755  
 A:Accession: B81040  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-298 <PEP>  
 A:Cross-references: GB:AE002530; GB:AE002098; MID:g7227054; PIDN:AAF42138.1; PID:g722  
 A:Experimental source: serogroup B, strain MC58  
 C:Genetics:  
 A:Gene: NMB1801

alignment\_scores:  
 Quality: 1577.00 Length: 298  
 Ratio: 5.292 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 99.664

alignment\_block:  
 US-09-303-518D-569 x B81040 ..

Align seg 1/1 to: B81040 from: 1 to: 298

```

1 ATGTTTCGTTTACATTCAGCGCTTTCCCTTGGCAACCCCATGCA 50
|||||
1 MetPheArgLeuGlnPheArgLeuPheProPheLeuArgThrAlaMetHis 17
51 CATCCGTGTGACCGCCCTGCTCAAAATGCTCTCCCTGCGCGCTTCT 100
|||||
17 sileLeuLeuThrAlaLeuLeuLysCysLeuSerLeuLeuProLeuSerc 34
101 GTCTGCACACGCTGGGAAACCGGCTCGACATCTGGCGTTTACCTTTA 150
|||||
34 yLeuHisThrLeuGlnLysnArgLeuGlnHisLeuAlaPheThrLeuLeu 50
151 AAGGAAGACCGCGCGCATGCTGCCCAATATGCTCAGCGAGCATGAA 200
|||||
51 LysGlnAspArgAlaArgIleValAlaAsnMetArgGlnAlaGlyMetAs 67
67 nProAspProLysThrValLysAlaValPheAlaGlnThrAlaLysGly 84

```

```

251 GTTTGGAAGTGGCCCGCGCTTTTTCAGAAACCGGAACATAGAAACA 300
    |||||||
84  LyeuGIleuAlaProAlaPhePheArgIysProGluAspIleGluThr 100
    |||||||
301 ATGTTCAAACGGGTACACGGCTGGGAACATGTCCAGCGCTTTGGACAA 350
    |||||||
101 MePheYsaIaValaHisGlyTrpGluHisValaGlnGlnAlaLeuAspIy 117
    |||||||
351 ACAGAAAGGGCTGTATTCATCAGCGCGACATCGGCACCTCGATTGG 400
    |||||||
117 SHISGluGlyLeuLeuPheHelleThrProHisIleGlySerTyrAspLeuG 134
    |||||||
401 GCGGAGCGTACATACAGCAGCAGCTTCCTCCGCTGACCGCAATGAC 450
    |||||||
134 LyeIaYrGlyTrpIleSerGlnGlnLeuProPheProLeuThrIaMetTyr 150
    |||||||
451 AAACCGCGCAAAATCAAGCGATAGCAAAATCATGACAGCGCGGAGGT 500
    |||||||
151 LysProProIyIleLysAlaIleAspIyIleMetGlnAlaGlyArgVa 167
    |||||||
501 TCGCGGCAAGGAAAGAAACCGCGCTTACAGATACAGAGGCTCAAAACA 550
    |||||||
167 LArgGlyLysGlyLysThrAlaProThrSerIleGlnGlyValIysGlnI 184
    |||||||
551 TCATCAAAAGCGCGCTGGGCGGAGCAACCATGCTCTCCGCGCAACAC 600
    |||||||
184 LelleLysAlaLeuArgSerGlyGlnAlaThrIleValLeuProAspHis 200
    |||||||
601 GTCCCGCTCCCTCAAGAAAGCGGAGGAGCGATGAGTGGATTCCTGG 650
    |||||||
201 ValProSerProGlnGlnGlyLysGlnGlyValITrpValAspPhePheI 217
    |||||||
651 CAACCTGCGCTATACATGACGCTGCGGCAAAATGCGACACGCTCAAG 700
    |||||||
217 LysProAlaIyTrpThrMetThrLeuAlaIaLysLeuAlaHisValIysG 234
    |||||||
701 GCGTAAACACCTGTTTTCGCTGCGAGCGCGCTGGGCGGCAACAGT 750
    |||||||
234 LyeValLysThrLeuPhePhePheCysGlyAlaGlyLeuProGlyLysGlnI 250
    |||||||
751 TTCGATTGGACATCCGCCCGCTCCAAAGGGAATGAAAGCGGACAAAGC 800
    |||||||
251 PheAspLeuHisIleArgProValGlnGlyLeuLeuAsnGlyAspIyAl 267
    |||||||
801 CCATGATGCGCGCGGTTCACACCGCAATGCCGAATATGATAGCGCGT 850
    |||||||
267 AhisAspIaIaValaPheAsnArgAsnAlaGlnIyTrpIleArgArgP 284
    |||||||
851 TTCGAGCGCAGTATCTTTTATGTACACCGCTACAAATGCGC 894
    |||||||
284 heProThrGlnTyrLeuPheMetTyrAsnArgTyrLysMetPro 298
    |||||||
seq_name: p1r2:C82691

```

```

seq_documentation_block:
  Lipid biosynthesis lauroyl acyltransferase XF1348 [imported] - Xylella fastidiosa (str
  C:Species: Xylella fastidiosa
  C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
  C:Accession: C82691
  R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
  Nature 406, 151-157, 2000
  A:title: The genome sequence of the plant pathogen Xylella fastidiosa.
  A:Reference number: A82515; MUID:20365717
  A:Note: for a complete list of authors see reference number A59328 below
  A:Accession: C82691
  A:status: preliminary
  A:Molecule type: DNA
  A:Residues: 1-316 <SIM>
  A:Cross-references: GB:AE003967; GB:AE003849; NID:g9106347; PIDN:AAFG4157.1; GSPDB:GN001
  A:Experimental source: strain 9a5c
  R:simpson, A.J.G.; Rehnach, F.C.; Arruda, P.; Abreu, F.A.; Agencio, M.; Alvarenga, R.; A
  Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H
  as Neto, E.; Docena, C.; El-Dorfy, H.; Facincanli, A.P.; Ferreira, A.J.S.

```

submitted to Genbank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Fr  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitejima, J.P.; Krieger, J.E.; Kuramae, E.B.; La  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.U.; Marques, M.V.; Martins  
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Mitaca, E.C.; Miyaki, C.  
 F.G.; Nunes, L.R.; Oliveira, M.A.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava  
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
 M.; Tsubako, M.H.; Vallada, H.; Van Slyys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.  
 A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: XF1348

alignment\_scores:  
 Quality: 263.50 Length: 286  
 Ratio: 1.532 Gaps: 9  
 Percent Similarity: 60.140 Percent Identity: 27.273

alignment\_block:

US-09-303-518D-569 x C82691 ..

Align seg 1/1 to: C82691 from: 1 to: 316

```

103 CTGCACACGCTGGGGAACCGGCTGGACATCTGGCGTTTACCTTTAA 152
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
10  LeuYrTrpValAlaSerLeuValGlyHisIleProTrpProPheLeuIy 26
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
153 GGAAGACCGCGCGCGCATCGCGC.....AATATGC 184
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
26  sArg...LeuSerAspLeuIleAlaIyTrpCysLeuAsnIaAla 42
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
185 GTGACGGA.....GGCATG 198
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
42  rGlnSerArgValThrGlnArgAsnLeuGluLeuThrTrpProGluLeu 58
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
199 AATCCGACCCCAAAAGC.....GTCAAGCGGTTTTCGGAACCGC 242
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
59  SerProGlnGlnHisThrArgLeuHisHisGlnIleIleYrSerThrVa 75
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
243 AAAAGCGGTTTGGAACTGCGCGCGGTTTTCAGAAACCGGAAGACA 292
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
75  LArgGlnAlaPheGluMetLeuHisIleTrpThrHisProGluThrGln 92
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
293 TAGAAACATGTTCAAAGCGGTACACCGCTGGGAACATGTGCGACGCT 342
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
92  snLeuAlaIaGlyLeuArgGluTyrHisGlyValaAspLeuTyrSerAlaIa 108
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
343 TTGGACAAACAGCAAGGCGCTGATTCATCAGCGCGACATCGCGAGCTA 392
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
109  LeuAlaArgAsnArgIyMetIleIleValaIaProHisPheIyAsnTr 125
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
393 CGATTGGCGGCGACGCTACATCAGCGACAGCTCCGTTCCGCGTACCG 442
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
125  pGluLeuLeuLysGlnTyrLeuSerAspLeuAla.....ProLeuThrL 140
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
443 CCATGTACAAACCGCGCAAAATCAAGCGATGACAAATCATGACAGGG 492
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
140  euMetIyTrpArgProAlaAsnSerAlaValaIleuAspIyPheLeuGln 155
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
493 GCGAGGCTTCGCGCAAGCAAAACCGCGCTACAGCATCAAGG. 540
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
156  ...ArgValaArgGlyLysAsnAsnValaHisGlnValaArgAlaGlnIy 171
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
541  ....GTCAACAAATCATCAAGCGCTGCGTTTCGCGGCAAGCAACCATG 586
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
171  laValaIaGlnHisLeuPheLysValaIleuLysAsnGlyLyrThrIleGlyI 188
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
587  TCCTGGCCGACACAGTCGCTCCCTCAAGAGGCGGGAAGCGCTATAG 636
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
188  leuLeuProAspGlnGlnPro.....LysIleGlyGlnIyValaIaPhe 201
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

637 GTGGATTCTTGGCAAAACCTGCTTACCATGACGCTGGCAAAATT 686
    : ||||| ||| ||||| ||||| ||||| ||||| |||||
202 AlarPhePheGlyIleGlyIleValIleuThrMetThrIleValIleArgIle 218
687 GCGACACGCTAAAGGCGTGAAGACCCGTTTCTGCTGGACAGGCTGC 736
    ||||| : ||||| : ||||| : ||||| : ||||| : |||||
218 uAlaGluArgThrSerAlaThrValIleuThrAlaTrpCysGluArgIle 235
737 CTGGCGGACAAGGTTTCGATTGTCACATCGCCGCTCCAAAGGGAATTG 786
    ||| ||||| : ||||| : ||||| : ||||| : ||||| : |||||
235 erProHISleuGluPheAlaIleuHISMetGluHISAlaSerProAlaVal 251
787 ... AACGGCCACAAGCCCATGATGCCGCGGTTCACCGCATGCCGA 833
    : ||| ||||| : ||||| : ||||| : ||||| : ||||| : |||||
252 AlaSPProSPProIleuIleAlaIleThrAlaIleuSPAlaGlyIleG 268
834 ATATTGGATACGCGCTTTCGACGACGATATCTGTTATGTACACCGCT 883
    | ||||| ||||| ||||| ||||| ||||| ||||| |||||
268 uArgIleAlaArgArgSPProThrGlnTrpGlnTrpThrTyrLysArg 285
884 ACAAAATG 891
    |||||
285 YrSerLeu 287

```

seq\_name: p1r2:H83642

```

seq_documentation_block:
probable 2-OH-lauroyltransferase PA0011 [Imported] - Pseudomonas aeruginosa (strain PA01
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83642
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bt
adman, S.; Tian, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337
A:Accession: H83642
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <STO>
A:Cross-references: GB:AE004441; GB:AE004091; NID:g9945828; PIDN:AA03401.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0011

```

```

alignment_scores:
Quality: 259.00 Length: 286
Ratio: 1.463 Gaps: 7
Percent Similarity: 61.888 Percent Identity: 25.874

```

alignment\_block:

US-09-303-518d-569 x H83642 ..

Align seg 1/1 to: H83642 from: 1 to: 295

```

55 CTGTGACGCGCTGCTCAAAATGCTCTCCCTGCTGCGCTTTCCTGCT 104
    ||||| : ||||| : ||||| : ||||| : ||||| : |||||
8 LeuValIleAlaGlyAlaLeuArgIleuHISMetThrIleValIleArgIle 24
105 GCAACGCTGGAAACCGGCTGACATCTGCGCTTACCTTTAAAG 154
    : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
24 IeGluValIleAlaGlyIleGlyIleTrpLeuMetTrpLysLeu.....P 39
155 AAGACCGCGCGGCGCATGCTGCCCATATGCGCGACGAGCATGAATCC 204
    : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
39 roAsnArgSerArgGluValValArgIleAsnLeuSerLysCysPhePro 55
205 GAC.....CCCAAAAGGTCAAAAGCGCTTTTGGCGGA 236
    : ||| ||| : ||| : ||| : ||| : ||| : |||
56 GluLeuSerGluThrGluLeuGluLysLeuValGlyGlnSerLeuMetAs 72

```

```

237 AACGGCAAAAGCGGTTTGGAACTTGGCCCCGCTTTTTCAGAAAACCGG 286
    : ||||| ||| ||||| ||||| ||||| ||||| |||||
72 PileGlyArgThrIleuThrIleuThrIleuThrIleuThrIleuThrIleu 89
287 AAGACATGAAACAATGTTCAAGCGGTACACGCGCTGGAAACATGTCAG 336
    ||| : ||||| : ||||| : ||||| : ||||| : |||||
89 IuLysSerLeuArgThrIleArgIleValGluGlyMetGluValIleGlu 105
337 CAGGCTTTGGCAAAACGAGGCGTGTATTCATTCACGCGCATCGG 386
    : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
106 GluAlaIleAlaIleSerGlySPLeuValIleGlyIleThrSerHISLeuG 122
122 YAsnTrpGluValIleAsnHISPheTyrCysSerTyrAlaLysProIle 138
387 CAGCTACGAT...TTGGCGGACGCTACATACGACGACGCTTCCGTTCC 433
    : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
434 CGCTACCGCCATGTACAAACCGCGAAATCAAAACGATGACAAATC 483
    : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
139 .....IlePheTyrArgProProLysLeuLysAlaValASPLeu 152
484 ATGACGCGCGGCGGTTTCGCGCAAGAAAACCGCGCTACACGAT 533
    : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
153 LeuLysLysGlnArgValGlnLeuGlnLysArgValAlaProSerThrPr 169
534 ACAAGGCGTCAAAACAATCATCAAGCCCTGCTGCGGCGCAAGACCA 583
    : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
169 oGluGlyIleLeuSerValIleLysGluValLysLysGlyGlyValG 186
584 TGCTGCTGCGGACGACGTCGCCCTCCCTCAAGAGCGGGAAGCGCTA 633
    : ||| ||||| ||||| ||||| ||||| ||||| |||||
186 LylIleProAlaSP.....ProGluProAlaArgThrAla...GlyLeu 199
634 TGGGTGGATTCTTCGCAAACTGCTTACCATGACGCTGGCGCAAA 683
    ||||| : ||||| : ||||| : ||||| : ||||| : |||||
200 PheValProTyrLeuGlyThrThrAlaLeuIleSerLysPheValProG 216
684 ATTGCGACACGTCMAAGGGGTGAACCTGTTTCTGCTCGAAGCC 733
    : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
216 nIleuLeuSerArgGlyLysAlaIleGlyAlaPhePheHISAlaValArg 233
734 TGCGTGGCGGACAAAGGTTTCGATTTGACATCCGCCCTCCCAAGGG 783
    ||||| : ||||| : ||||| : ||||| : ||||| : |||||
233 eutroAspLysSerGlyTyrLysValIleLeuGluAlaIleProAlaSP 249
784 TTG...AACGGCGCAAAAGCCCATGATGCCGCGGTTCACCGCAATGC 830
    : ||| : ||||| : ||||| : ||||| : ||||| : |||||
250 MetTyrASPlySPLeuGluValSerValAlaAlaMetSerArgGlu 266
831 CGAATATTGATACGCGCTTTCGACGACGATCTGTTATGTACACCC 880
    : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
266 uAlaLysTyrValArgAlaIleTyrProSerGlnTyrMetTrpSerMetLys 283
881 GCTACAAA 888
    |||||
283 rPheLys 285

```

seq\_name: p1r2:D83239

```

seq_documentation_block:
probable lauroyl acyltransferase PA3242 [Imported] - Pseudomonas aeruginosa (strain P
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: D83239
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bt
adman, S.; Tian, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: D83239
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 <STO>

```

A:Cross-references: GB:AE004747; GB:AE004091; NID:99949362; PIDN:ANG06630.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA3242

alignment\_scores:  
 Quality: 228.50 Length: 281  
 Ratio: 1.336 Gaps: 6  
 Percent Similarity: 60.854 Percent Identity: 26.335

alignment\_block:  
 US-09-303-518D-569 x D83239 ..

Align seg 1/1 to: D83239 from: 1 to: 312

```

67 CTGCTCAATGCTCTCCCTGCTGCGCTTTCTGTGTCACACGCTGG 116
   |||||      |||||      |||||      |||||
24 LeuLeuTrpLeuValValGlnLeuProTyrProValLeuLeuMetLeuG 40
   |||||      |||||      |||||      |||||
117 AACCGGCTCGACATCTGGCTTTTACTTTTAAAGGAGACCGCGCG 166
   |||||      |||||      |||||      |||||
40 YArgGlyLeuGlyAlaLeuMetTyrArgLeuValGlySerArgArgGlu 57
   |||||      |||||      |||||      |||||
167 GCATGCTGCCAATATGCTCAGCAGGATCATCCGACCC... 210
   |||||      |||||      |||||      |||||
57 LeuAlaAlaArgAsnLeuGlnLeuGlyPheProGlnLysSerProAlaGlu 73
   |||||      |||||      |||||      |||||
211 .....AAACGGTCAAACCGCTTTTGGGAAACGGCAAAAGCGGTTT 254
   |||||      |||||      |||||      |||||
74 ArgGluArgLeuLeuLysGlnAsnPheAlaSerSerGlyIleAlaPhePh 90
   |||||      |||||      |||||      |||||
255 GGACTTGGCCCGCGCTTTTCAGAAACCGGAGACATAGAACATCT 304
   |||||      |||||      |||||      |||||
90 eGluMetAlaMetSerTrpTrp...TrpProGlyAlaArgLeuAlaArgL 106
   |||||      |||||      |||||      |||||
305 TCAAACGGGACACGCTGCGGAAACATGTCAGAGGCTTTGGCAACAC 354
   |||||      |||||      |||||      |||||
106 euuAlaHisIleGlnGlyLeuGlnHisLeuArgGlnAlaGlnAlaGlnGly 122
   |||||      |||||      |||||      |||||
355 GAAGGCTGCTATTTCATCAGCGCGCATCGCGAGCTACGATTTGGCGG 404
   |||||      |||||      |||||      |||||
123 GlnGlyValIleLeuMetAlaLeuHisPheThrThrLeuGlnIleLeuAl 139
   |||||      |||||      |||||      |||||
405 AGCGTACATCAGCCAGACGCTTCCTCCCGCTGACCGCCATGTACAAAC 454
   |||||      |||||      |||||      |||||
139 AlaLeuLeuGlnGlyGln.....ValHisThrIleAspGlyMetTyrArg 154
   |||||      |||||      |||||      |||||
455 CGCGGAAATTCAAAGGATAGCAAAATCATGCGCGGCGAGGTTGCG 504
   |||||      |||||      |||||      |||||
154 LuHisAspAsnProValPheAspTyrValGlnArgArgGlnArgGlnArg 170
   |||||      |||||      |||||      |||||
505 GGCAAGGAAAGAAACCGCGCTACACATACAGGGGTGCAACAAATCAT 554
   |||||      |||||      |||||      |||||
171 HisAsnLeuAspAlaThrAlaIleGlnArgGlnAspValArgSerMetLe 187
   |||||      |||||      |||||      |||||
555 CAAGCGCTGCGCTGGCGGAGCAACCATGCTCGCCCGACACAGCTCC 604
   |||||      |||||      |||||      |||||
187 uLysValLeuArgGlyGlyArgAlaIleTrpTyrAlaProAsp..... 201
   |||||      |||||      |||||      |||||
605 CCTCCCTCAAGAGAGCGGG.....GAAGCGGTATGGGATTTCTTC 648
   |||||      |||||      |||||      |||||
202 .....GlnAspTyrGlyAlaLysGlnSerLeuPheValProLeuPhe 215
   |||||      |||||      |||||      |||||
649 GGCAGAACCTGCTTATACATGACGCTGGCGGCAAAATTTGGACACGTC 698
   |||||      |||||      |||||      |||||
216 GlyIleProAlaAlaThrValThrAlaThrThrLysPheAlaArgLeuG 232
   |||||      |||||      |||||      |||||
699 AGCGGTGAGAAACCTGTTTCTGCTGCGAGACCGCTGCGCGAGCAAG 748
   |||||      |||||      |||||      |||||
232 YArgAlaArgValLeuProPheThrGlnSerArgLeuAlaAspGlySer 249
   |||||      |||||      |||||      |||||
749 GTTTCGATTTCGACATCCGCCCGTCCAGAGGGAATTGAACGGCGACAAA 798
   |||||      |||||      |||||      |||||

```

seq\_name: pir2:S16888

seq\_documentation\_block:

lipid A biosynthesis lauroyl acyltransferase (EC 2.3.1.-) - Escherichia coli  
 N:Alternate names: htrb protein  
 C:Species: Escherichia coli  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 08-Oct-1999  
 C:Accession: S16888; A42290; C64848  
 R:Karrow, M.; Georgopoulos, C.  
 Mol. Microbiol. 5: 2285-2292, 1991  
 A:Title: Sequencing, mutational analysis, and transcriptional regulation of the Esche  
 A:Reference number: S16888; M01D:92114808  
 A:Accession: S16888  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-306 <KAR>  
 A:Cross-references: EMBL:X61000; NID:948950; PIDN:CAA43317.1; PID:948957  
 A:Experimental source: strain K-12  
 R:Ueguchi, C.; Ito, K.  
 J. Bacteriol. 174, 1454-1461, 1992  
 A:Title: Multicopy suppression: an approach to understanding intracellular function in  
 A:Reference number: A42290; M01D:92165719  
 A:Accession: A42290  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 200-306 <UEG>  
 A:Cross-references: EMBL:X59939; NID:942028  
 A:Experimental source: strain K-12  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
 A:Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; M01D:97426617  
 A:Accession: C64848  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-306 <BLAT>  
 A:Cross-references: GB:AE000206; GB:U00096; NID:91787282; PIDN:MAC74138.1; PID:917872  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genetics:  
 A:Gene: htrb  
 C:Function:  
 A:Description: acylates the intermediate (kdo)2-lipid IVA  
 A:Pathway: lipopolysaccharide biosynthesis  
 A:Note: required for bacterial viability above 33 degrees C in rich media  
 C:Keywords: acyltransferase; lipopolysaccharide biosynthesis; transmembrane protein  
 F:21-37/Domain: transmembrane #status predicted <TMM>

alignment\_scores:

Quality: 215.00 Length: 274  
 Ratio: 1.303 Gaps: 9  
 Percent Similarity: 60.219 Percent Identity: 29.197

alignment\_block:

US-09-303-518D-569 x S16888 ..

Align seg 1/1 to: S16888 from: 1 to: 306

```

88 CTGCGCTTTCCTGTCGACACGCTGGGAAACCGCTCGACATCTGCG 137
   |||||      |||||      |||||      |||||
32 LeuProTyrProValIleTyrArgLeuGlnGlyGlyGlnGlyLysLeuAl 48
   |||||      |||||      |||||      |||||

```







Percent Similarity: 57.092 Percent Identity: 29.078

alignment\_block:  
US-09-303-518D-569 x A10441 ..

Align seg 1/1 to: A10441 from: 1 to: 306

```

67 CTGCTCAATAGCTCTCCCTGCTGCGCTTCTGCTGCACACGCTGG 116
   ::::: ||| ||||| |||||
25 VALLEupheLeuValGlnLeuProtyrProLeuValAsnLysLeuG1 41
   ::::: ||| ||||| |||||
117 AACCGGCTCGACATCTGGCGCTTTTACCTTTAAAGAAAGCGCGCGC 166
   ::::: ||| ||||| |||||
41 yAltrLeuGlyArgThrSerMetArgPheLeuLys.....ArgArgV 56
167 GCATGGTGGCAATATGGCTCAGGACGACATGAATCCGAC..... 207
56 aISerLeuAlaArgAsnLeuGluLeuLysPheProAspMetCAspLys 72
208 .....CCCAAAACGGTCAAAACCGCTTTTGGCGAAACGGCAAAAG 248
73 GlnValLeuGlnGlnThrLeuLeuGlnPheGlnSerLeuGlyMetG1 89
249 CGGTTTGGCACTTGGCCCGCTTTTTCAGAAACCGCAACGACATAGAA 298
89 yLeuLeuGlnThrGlyMetAlaTrpPheTrpSerAspThrArgLysGln 106
299 CAATGTTCAAGCGGTACACGCGTGGGAACATGTCACAGCGCTTGGAC 348
106 rGTTrpPhe...SerValSerGlyLeuGlnAsnLeuLysArgAlaGln 121
349 AAACACGAAGGGCTCTATTTCATCAGCGCGCACATGGCGACCTAGATT 398
122 GlyLysArgGlyValLeuValLeuGlyValHisPheMetSerLeuGlu 138
399 GGGCGGACGCTCATCAGCAGCAGCTTCCGCTCCGCTGACCGCGCATGT 448
138 uGlyGlyArgValMetGlyGlnCysGln.....ProMetLeuAlaMet 153
449 ACAAAACCGCGCAAAATCAAAAGCATAGACAAATCATGACGCGGCGAG 498
153 yArgProHisAsnAsnLysValMetGluLeuValGlnThrTrpLysArg 169
499 GTTCCGGCGCAAGAAACCGCGCTTACCATCAGCATACAAAGGGCTACA 548
170 MetArg.....SerAsnLysAlaMetLeuSerArgArgAspLeuArgG1 184
549 AATCATCAAGCGCTGCGCTGGCGGAAGCAACATCGCTCGCGCGCAC 598
184 yMetValArgAlaLeuLysGlnGlyGlnAlaValTrpPheAlaProAspG 201
599 ACGTCCCTCCCTCCCAAGAAAGCGGGGAAGCGTATGGTGATTTCTTC 648
201 InAspTrpLysProArg.....GlySerValPheAlaProLeuPhe 214
649 GGC...AAACSTGCTATACATGACGCTGGCGGCAAAATGGCACAGCT 695
215 AlaValGlnHisAlaAlaThrThrSerGlyThrPheMetLeuAlaArg 231
696 CAAGGCGTGAAGAACCTGTTTCTGCTGGCAAGCGCTGCGCGGAC 745
231 uAlaLysProAlaLeuLeuProLeuValLeuLeuArgLysLysGlnGly 248
746 AAGTTTTCATTTGCATCCGCGCTCCCAAGGGAATTGAAC...GGC 792
248 rGgGlyTrpAspLeuLeuLeuGlnProAlaLeuGlnAspTrpProLeuAsp 264
793 GACAAAGCCCATGATCCGCGCTGTTCACACGCAATGCCGAATATGGAT 842
265 AspGlnLeuAlaAlaLaserGlyMetAsnLysValLeuGlnGlnL1 281
843 ACGCGTTTTCGACGACGATCTGTTATATGATACAAACGCTTACAA 888
   ||| ||| ||||| ::::: |||

```

281 emetArgAlaProGluGlnTyLeuTrpLeuHisArgArgPheLys 296  
seq\_name: p1r2:D81085

seq\_documentation\_block:  
HitB/MSDB family protein NMB1418 [imported] - Neisseria meningitidis (strain MC58 ser C:Species: Neisseria meningitidis  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
R:Accession: D81085  
C:Retellin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. et al.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizzi, M. Science 287, 1809-1815, 2000  
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. A:Reference number: A81000; MUID:20175755  
A:Accession: D81085  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-289 <TEXT>  
A:Cross-references: GB:AE002491; GB:AE002098; NID:g7226655; PIDN:AAF41779.1; PID:g722 A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NMB1418

alignment\_scores:  
Quality: 204.50 Length: 286  
Ratio: 1.225 Gaps: 11  
Percent Similarity: 58.392 Percent Identity: 26.923

alignment\_block:  
US-09-303-518D-569 x D81085 ..

Align seg 1/1 to: D81085 from: 1 to: 289

```

67 CTGCTCAATAGCTCTCCCTGCTGCGCTTCTGCTGCACACGCTGG 116
   ::::: ||| ||||| |||||
7 ValLeuTrpValLeuGlnPheLeuProAlaLeuLeuHisLysL1 23
117 AAACCGGCTCGACATCTGCGCTTTTACCTTTAAAGAAAGCGCGCGC 166
   ::::: ||| ||||| |||||
23 aAspLeuThrGlyLeuLeuAla...TyLeuLeuValLysProArg 39
167 GCATC.....GTGCCAATATGCGTACGCGACGATGAATCC 204
39 rGlyGlyGlnLeuAsnLeuAlaLysCysPheSerGlnTrpSerGln 55
205 GACCCCAAAACGTC...AAAGCGTTTTCGCGAAACGGCAAAAGCGG 251
56 LysArgLysThrValLeuLysGlnHisPheLysHisMetAlaLysLeu 72
252 TTGGAACCTTGGCCCGCTTTTTCAGAAACCGGAACATAGATGAACA 301
72 tLeuGlnTyrcLysLeuTrpTrpTrpArgProAlaGlyArgLeuSer 89
302 TGTTCAAAGCGGTACACGCGTGGGAACATGTCACAGCTTTGCAACA 351
89 euValArg...TyArgAsnLysHisTyLeuAspAspAlaLeuAla 104
352 CACGAAGGCTGCTATTCATCAGCGCGCACATGGCAGATAGATTTGG 401
105 GlyLysValAlaLeuLeuLeuTyTrpHisPheThrAlaPheGlnMet 121
402 CGAGCGTATCATCAGCAGCGTCCGCTTCCGCTGACCGCATGTACA 451
121 aValTyAlaLeuAsnGlnAspLeu.....ProLeuLieserMetLys 136
452 AACCGCGCAAAATCAAGCATAGACAAATCATGACGCGCGAGGTT 501
136 erHisGlnLysAsnLysLeuLeuAspGlnLeuLeuLysGlyArgAsn 152
502 CGCGCGCAAGAAACCGCGCTACACGATACAAAGGGTCAAAACAT 551

```



```

153 ArgTyrHisAsnValPheLeuIleGlyArgThrGlyLeuArgAlaIle 169
552 CATCAAGCCCTGGCTTCGGGCGAAGCAAC...ATCGTCCTCCGACCC 598
169 uValLysGlnPheArgLysSerSerAlaProPheLeuTyrLeuProAsp 185
599 ACGTCCCTCCCTCAAGAGCGGG...GAAGGGGTATGGGTGAT 642
186 .....GlnAspPheGlyArgAsnAspSerValPheValAsp 197
643 TTCTTGGCAACCTGCTATACCATGACGCGCGCAAAATGTGCACA 692
198 PhePheGlyIleGlnThrAlaThrIleThrGlyLeuSerArgIleAlaAl 214
693 CGTCAAGCGGTGAAGAACCCCTTTTCTGCGCCAGACCCCTGCT... 738
214 aLeuAlaAsnAlaLysValIle.....ProAlaIleProValA 227
739 .....GGCGACACAGTTTCGATTTCACATCCGCCCTCCACAGGGGA 783
227 TGGTAlaAspAsnThrValThrLeuHisPheTyrProAlaThrLysSer 243
784 TTGAACGGCGACAAAGCCCAT...GATCGCGCGGTTCACCGCAATGC 830
244 PheProGlyGlnAspAlaLysAlaAspAlaGlnArgMetAsnArgPheI 260
831 CGAATATTGATACGCGTTTCCGACGCGCATCTCTTATGTCACAC 880
260 eGlnAspArgValArgGlnHisProGlnGlnTyrPheThrLeuHisLysA 277
881 GCTACAA 888
277 rGpHelys 279

seq_name: pir2:B81857

seq_documentation_block:
probable acetyltransferase NMA1630 [imported] - Neisseria meningitidis (strain Z2491 ser
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_rev1sion 05-May-2000 #text_change 02-Feb-2001
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556
A:Accession: B81857
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <PAR>
A:Cross-references: GB:AL162756; GB:AL157959; NID:97380091; PIDN:CA84858.1; PID:9738027
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1630

alignment_scores:
Quality: 203.50 Length: 285
Ratio: 1.219 Gaps: 9
Percent Similarity: 58.596 Percent Identity: 25.965

alignment_block:
US-09-303-518D-569 x B81857 ..

Align seg 1/1 to: B81857 from: 1 to: 289

```

```

23 aAspLeuThrGlyLeuLeuAlaTyrLeuLeuValLysProArgArgArgI 40
167 GCATCTGCGCCATATATGCGTCAGCAGCATGATCCGCCCAAAAGC 216
40 IeGlyGlnIleAsnLeuAlaLysCysPheProGlnTyrPaspGlyLys 56
217 GTCAAGCCGTT.....TTTGGGAAACGGCAAAAGCGGTTT 254
57 ArgLysThrValLeuLysGlnHisPheLysHisMetAlaLysLeuMet 73
255 GGAACCTGGCCCCCGTTTTCAGAAAACCGGAACACATGAAACATGT 304
73 uGlnTyrGlyLeuTyrTrpTyrAlaProAlaGlyArgLeuLysSerLeu 90
305 TCAAGCGGTACACAGCGCTGGACATGTGCAGAGCTTGTGCAAAAC 354
90 alArg...TyrArgAsnLysHisTyrLeuAspAspAlaLeuAlaGly 105
355 GAAGGCTGCTATTCATCAGCGCGCACATGCGACGCTAGCATTTGGCG 404
106 GlnLysValIleIleLeuTyrProHisPheThrAlaPheGlnMetAla 122
405 ACGCTACATCAGCGACGCTTCGCTCCGCTGACCGCATGACAAAC 454
122 ITrAlaLeuAsnGlnAspVal.....ProLeuIleSerMetLysSer 137
455 CGCCGAAATCAAGCGATAGACAAATCATCAGCGCGGCGAGGTTTCG 504
137 IsglnLysAsnLysIleLeuAspGlnGlnIleLeuLysGlyArgAsnArg 153
505 GCGAAAGAAAACCGCGCTACCAGATACAGGGGTCAACAAATCAT 554
154 TyrHisAsnValPheLeuIleGlyArgThrGlnGlyLeuArgAlaLeu 170
555 CAAGCCCTGCGTTTCGGCGCAACCAAC...ATCGTCCTCCGCGACAG 601
170 LlysGlnPheArgLysSerSerAlaProPheLeuTyrLeuProAsp... 185
602 TCCCTCCCTCAAGAGCGGG.....GAGGCTATGGGTGATTTTC 645
186 .....GlnAspPheGlyArgAsnAspSerValPheValAspPhe 198
646 TTGCGCAACCTGCTATACCATGACGCGCGCGCAAAATGTGCACAG 695
199 PheGlyIleArgThrAlaThrIleThrGlyLeuSerArgIleAlaIle 215
696 CAAGCGCGTGAAGAACCCCTGTTTCTGTCGCAACCGCTGCT... 738
215 uAlaAsnAlaLysValIle.....ProAlaIleProValArg 228
739 .....GGCGACACAGTTTCGATTTCACATCCGCCCTCCACAGGGAAT 786
228 IuAlaAspAsnThrValThrLeuHisPheTyrProAlaThrLysSerPhe 244
787 AACGGCGACAAAGCCCAT...GATCGCGCGGTTCACCGCAATGCCGA 833
245 ProSerGlnAspAlaGlnAlaAspAlaGlnArgMetAsnArgPheIle 261
834 ATATTGGATACCGCGTTTCCGACGCGATGCTGTTATGTCACACCG 883
261 uGlnArgValArgGlnHisProGlnGlnTyrPheThrLeuHisLysArg 278
884 ACADA 888
278 helys 279

seq_name: pir2:AD0807

seq_documentation_block:
probable acyltransferase (EC 2.3.1.-) [imported] - Salmonella enterica subsp. enteric
C:Species: Salmonella enterica subsp. enterica serovar Typh
A>Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_rev1sion 09-Nov-2001 #text_change 27-Nov-2001

```

C:Accession: AD0807  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmons, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov A:Reference number: AB0502; PMID:11677608  
 A:Accession: AD0807  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-306 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD07638.1; PID:q16503625; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STY2639  
 C:Keywords: acyltransferase

alignment\_scores:  
 Quality: 199.50 Length: 296  
 Ratio: 1.255 Gaps: 9  
 Percent Similarity: 53.716 Percent Identity: 26.014

alignment\_block:  
 US-09-303-518D-569 x AD0807 ..

Align seg 1/1 to: AD0807 from: 1 to: 306

```

67 CTGCTCAATGCTCTCTGCTGCTGCTCTCTCTGTCACAGCGTGG 116
   ::::: ||| ||||| ||||| |||||
25 ValLeuTrpLeuValGlnLeuProTyrProValLeuArgPheLeuG 41
117 AAACGGCGTGGACATCTGGCGTTTACCTTTAAAG..... 153
   ::::: ||| ||||| ||||| |||||
41 YThrArgThrGlyLysLeuAlaArgProPheLeuLysArgArgLys 58
154 .....GAGAC 159
58 LeuGlnLysnLleGluLeuGlyPheProThrLeuSerArgLysL 74
160 CGCGGCGCATGCTGCC...AATATCGCTCAGCGCAGCATGATCCGA 206
   ||| ::::: ||| ||||| |||||
75 ArgGlyLysLeuLleAlaGlnLysnLysSerLeuGlyMet..... 88
207 CCCCAAAAGGTCAAAAGCGTTTGGGAAACGCGCAAAAGCGGTTGG 256
89 .....AlaLeuLeuG 92
257 AACTGCCCCCGCTTTTTCAGAAAACGGAACATAGTAAACAATGTC 306
   || ::::: ||| ::::: |||
92 LThrGlyMetAlaTrpPheTrpProAspSerArgValArgLysTrpPhe 108
307 AAAGCGGTACAGCGTGGGAAATGTCAGCAGCGTTTGGCAAAACAGA 356
   ||| ||| ::::: |||
109 Asp...ValAspGlyLeuAspAsnLeuThrArgAlaGlnAlaGlnAsn 124
357 AGGCGGTATTCATCAGCAGCGCATATGTCAGTACGATTTGGCGGAG 406
   ||| ::::: ||| ||||| |||||
124 GgLYValMetValGlyValHisPheMetSerLeuGluLeuGlyGly 141
407 GCTACATCAGCAGCGTCCGCTCCGCTGACCGCCATGTACAAACCG 456
   || ::::: ||| ||||| |||||
141 rGValMetGly.....LeuGlyGlnProMetMetAlaThrArgPro 155
457 CGGAATATCAAGCGTATGACAAATCATGACGCGCGAGGAGGTTGCG.. 504
   ::::: ||| ||||| |||||
156 HisAsnAsnProLeuMetGluTrpValGlnThrArgGlyArgMetArg 172
505 ..GGCAAGGAAAAACCGCGCTACACGATACAAAGGGTCAAAACAATCA 553
   ::::: ||| ||||| |||||
172 rAsnLysAlaMetLleGlyArgAsnAsnLeuArgGly.....LleV 186
554 TCAAAGCCCTGCGTTGGGCGAAAGCAACATGCTCTGCCGACACAGTC 603

```

```

:: ||||| ::||| ::|||
186 AlGValAlaLeuLysLysGlyAlaValAlaTrpPheAlaProAspLys 202
604 CCCTCCCTCAAGAAAGCGGCGGAGCGGATGCGGATTTCTGGG... 651
   ::||| ::||| ::|||
203 TyrGlyProLysLysSer.....PheAlaProPheAlaVal 216
652 ..AAACCTGCTATACCATGACGCTGGCGCAAAATTTGACACGTCAAAG 700
   ::||| ::||| ::|||
216 lGluAsnValAlaThrThrAsnGlyThrTyrValLeuSerArgLeuS 233
233 lYAlaAlaMetLeuThrValThrMetValArgLysSerAspAsnSer 249
701 GCGTGAACCCCTGTTTCTGCTGCGAACGCTGCTGGCGCAACAGT 750
   ||| ::||| ::|||
751 TTGCGATTGACATCCCGCCCGCGCAAGGGAATTGACGCGCGT 800
   ::||| ::||| ::|||
250 TyrArgLeuTyrThrThrProGluMetGluGlyTyrProAlaAsp 266
801 CCATGATGCCCGCGTTCACCGCAATGCGAATTTGGATACGCGCT 850
   ::||| ::||| ::|||
266 nGlnAlaAlaAlaTyrMetAsnLysLleGluLysGluLleMetArg 283
851 TTCCGACGCAATGCTTTATGTATGATACACCGCTACAA 888
   ||| ||||| ||||| ::|||
283 lArProGluGlnTyrLeuTrpLleHisArgArgPheLys 295

```

seq\_name: pIR2:E82847

seq\_documentation\_block:  
 Lipid A biosynthesis lauroyl acyltransferase XF0104 [imported] - *Xylella fastidiosa* (*C:Species: Xylella fastidiosa*  
 C:Date: 18-Aug-2000 #sequence revision 20-Aug-2000 #text change 20-Aug-2000  
 C:Accession: E82847  
 R:Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Seq  
 Nature 406, 151-157, 2000  
 A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
 A:Reference number: AB2515; MUID:20365717  
 A:Note: for a complete list of authors see reference number A59328 below  
 A:Accession: E82847  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-279 <SIN>  
 A:Cross-references: GB:AE003864; GB:AE003849; NID:q9104879; PIDN:AF82917.1; GSPDB:GN  
 R:Experimental source: strain 945c  
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre  
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr  
 J.D.; Junqueira, M.L.; Kempf, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E.; La  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins  
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.  
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmiter,  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa  
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Verjovski-Almeida, S.; Vettore, A.L.  
 M.; Tsunaho, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.  
 A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: XF0104

alignment\_scores:  
 Quality: 196.50 Length: 275  
 Ratio: 1.244 Gaps: 10  
 Percent Similarity: 57.455 Percent Identity: 27.273

alignment\_block:  
 US-09-303-518D-569 x E82847 ..

Align seg 1/1 to: E82847 from: 1 to: 279

```

88 CTGCCGCTTTCCTGTCACACGCTGGGAACGCGCTCGACATCTGCG 137
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
7  LeuProThrIleuGlnArgArgLeuGlyArgValIleGlyAlaMetAl 23
138 GTTTTACCTTTTAAAGAAACCGCGCGCATCGTCCAAATAGCGTC 187
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
23 AlaMetArgLeuValGlyThrArgArgCysAlaAlaGlnValAsnLeuArgL 40
188 AGCGAGCGCATGAATCCCGAC.....CCCAAAACGGTCGCAAGCC 225
   : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
40 eucysPheProThrIleHisAspGlnAlaThrArgIleGlyLeuHisGln 56
226 GTTTTTCGCGGAACGCGCAAGCGGTTTGGAACTTGCCTCCCGCTTTT 275
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
57 AsnPheAspAlaLeuGlyValGlyLeuPheGlnPheAlaArgAlaThrPr 73
276 CAGAAAAACCGAAGACATAGAAACAATGTTCAAGCGGTACAC.....G 319
   : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
73 P.....GlySerIleAspValIleArgAlaGlyIleHisValGluG 87
320 GCTGGGACATGTGTCAGACGCTTGGCAACACGAGGCGGTATTC 369
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
87 ILeuGlnHisLeuHisGlnLeuGlnArgGluValArgGlyValLeuLeu 103
370 ATGACCGCGCACTCGGACGCTAGTTGGCGGACGCTACATCGCA 419
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
104 ValSerGlyHisPheMetThrLeuGlnIleCysGlyArgLeuLeuLeu 120
420 GCAGCTCCGTTCCCGCTGACCGGCTGTACAAACGCGCAAAATCAAG 469
   : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
120 pTyVal.....ProLeuAlaGlyMetGlyThrArgHisArgAsnProV 135
470 CGATGACAAAAATCATGACGCGGCGAGGTTGCGGCAAGAAAAACC 519
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
135 allLeuGlnThrAlaIleLysArgGlyArgLeuArg.....TyrAlaThr 149
520 GCGCCATACGACATACAGGGGTCAAAACATCATCAAGCCCTCGCTTC 569
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
150 GlnMetPheAlaAsnGlnIleuLysArgAlaThrIleLysHisLeuLysAr 166
570 GGGCGAAGCAACCATGTCCTGCTGCGGACGCTCCCTCCCTCAAGAG 619
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
166 gGlyGlyPheLeuThrPylAlaProAsp.....GlnAspMetArgG 180
620 GCGGGGAAGGCTATGGGTGATTTCTTGGCAACGCTGCTATACATG 669
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
180 ILeuGlnIleThrValPheAlaProPhePheGlyMetProAlaIleThrIle 196
670 ACGCTGGCGCAAAATGCGACACGTCGCAAGGCGTGAACCCCTG...TT 716
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
197 ThrAlaThrHisGlnLeuAlaIleArgLeuThrGlyCysAlaValAlaProTy 213
717 TTTTCTGCTGCGAAGCGCTGCTGCGGACGCAAGGTTTCGATTTCACATCC 766
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
213 rPhe.....HisArgArgGlnGlyGly...AsnThrValLeuLysIleA 227
767 GCCCGCTCCAGGGAATG...AAGCGCGCAAAACCCCATGATGCGCGC 813
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
227 IArgProLeuAlaAspPheProSerGluAspLeuAlaAlaAspThrAla 243
814 GTTTCACCGCAATGCGCAATATGTCATACCGGCTTTCGAGGCGCA 863
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
244 ArgIleAsnGlnValIleGlyThrMetIleAspGlnAlaIleProAspGlnTy 260
864 TCTGTATTGTACACCGCTCAAA 888
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
260 rLeuThrIleHisArgArgPheLys 268

```

seq\_name: p1r2:B91036

seq\_documentation\_block:  
 probable heat shock protein [imported] - Escherichia coli (strain 0157:H7, substrain RIM  
 C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
 C;Accession: B91036  
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
 gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and g  
 A;Reference number: A9629; MUID:21156231; PMID:11258796  
 A;Accession: B91036  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-328 <NAV>  
 A;Cross-references: GB:BA000007; PID:BA036681.1; PID:g13362728; GSPDB:GN00154  
 A;Experimental source: strain 0157:H7, substrain RIMD 0509952  
 C;Genetics:  
 A;Gene: ECS3258

alignment\_scores:  
 Quality: 189.50 Length: 294  
 Ratio: 1.199 Gaps: 12  
 Percent Similarity: 53.741 Percent Identity: 26.190

alignment\_block:

US-09-303-518D-569 x B91036 ..

Align seg 1/1 to: B91036 from: 1 to: 328

```

88 CTGCCGCTTTCCTGTCACACGCTGGGAACGCGCTCGACATCTGCG 137
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
54 LeuProThrIleuGlnArgArgLeuGlyArgValIleGlyAlaMetAl 70
138 GTTTTACCTTTTAAAGAAACCGCGCGCATCGTCCAAATAGCGTC 187
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
70 AArgProPheLeuHisAspGlnAlaThrArgIleGlyLeuHisGln 87
154 .....GAGACCGCGCGCGCATCGTCC... 177
87 eucysPheProThrIleHisAspGlnAlaThrArgIleGlyLeuHisGln 103
178 AATTTGCGTCAAGGCGCATGATCCGACCCCAAAAGGTCGCAAGCGGT 227
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
104 AsnPheArgSerLeuGlyMet..... 110
228 TTTTTCGCGAAGCGGCGGTTTGAACCTTCCCGCGCTTTTCA 277
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
111 .....AlaLeuValGlnThrGlyMetAlaThrPhe 121
278 GAAACCGGAAGACATGAAACAATGTTCAAGCGGTACGCGCTGGAA 327
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
121 rProAspSerArgValArgLysThrPheAsp...ValGlnGlyLeuAsp 136
328 CATTTGACGACGCTTTGACAAACGCAAGGCTGCTATTCATCAAGCGC 377
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
137 AsnLeuLysArgAlaGlnMetGlnAsnArgGlyValMetValValGlyA 153
378 GCACATGCGGACGCTACGATTTGGCGGACGCTACATGACGACACGCTTC 427
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
153 HisPheMetSerLeuGlnIleuGlyArgValMetGly.....LeuC 168
428 GCTTCCCGCTGACCGCATGATACAAACCGCGCAAAATCAAGGATAGAC 477
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
168 yGlnProMetAlaThrThrArgProHisAsnGlnLeuMetGlu 184
478 AAATCATGTCAGCGGCGGCGGCTTCC...GCCAAGGAAAAACCGCGC 524
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
185 TrpValGlnThrArgLysArgMetArgSerAsnLysAlaMetIleLysAr 201
525 TACACGATACAAAGGCGTCAAACAATCAACAAGCCCTGCTGCGGCG 574
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||
201 gAsnAsnLeuArgGly.....IleValGlyAlaLeuLysLysGlyG 215
575 AACCAACCATGTCCTGCCCGACACGCTCCCTCCCTCAAGAGCGCG 624
   ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||

```



A: Molecule type: DNA  
 A: Residues: 1-328 <STO>  
 A: Cross-references: GB:AE005174; NID:912516752; PIDN:AG57504.1; GSPDB:GN00145; UWGP:236  
 A: Experimental source: strain 0157:H7, substrain EDL933  
 C: Genetics:  
 A: Gene: ddg

alignment\_scores:  
 Quality: 189.50 Length: 294  
 Ratio: 1.199 Gaps: 12  
 Percent Similarity: 53.741 Percent Identity: 26.190

alignment\_block:

US-09-303-518D-569 x D85880 ..

Align seg 1/1 to: D85880 from: 1 to: 328

```

88 CTGGCGCTTCTGTCGACACGCTGGGAAACGGCTGCGACATCTGGC 137
   |||||
54 LeuProTyrProValLeuLysPheLeuGlyThrArgIleGlyAlaMetAl 70
   |||||
138 GTTTTACCTTTTAAAG..... 153
   |||||
70 AlaArgProPheLeuLysArgArgIleSerIleAlaArgLysAsnLeuGlu 87
   |||||
154 .....GAAGACGGCGCGCGCATCGTCGCC... 177
   |||||
87 eucysPheProGlnHisSerAlaGlnIleValArgLysMetIleAlaGlu 103
   |||||
178 AATATGCGCGACGAGCGATGAATCCGACCCCAAAACGGCTCAAGCCGT 227
   |||||
104 AsnPheArgSerLeuGlyMet..... 110
   |||||
228 TTTTCCGGAACGCAAAAGCGGCTTTGGAATTCGCCCGCTTTTCA 277
   |||||
111 .....AlaLeuValGluThrGlyMetAlaIlePheT 121
   |||||
278 GAAACCGGAGACATAGTAACAAATGTTCAACCGCTACACGGCTGGAA 327
   |||||
121 rPrProAspSerArgValArgLysTrpPheAsp...ValGluGlyLeuAsp 136
   |||||
328 CATGTGACGAGCGCTTTGACAAACACGAGCGCTATTCATCACGCC 377
   |||||
137 AsnLeuLysArgAlaGlnMetClnAsnArgLysValMetValIleGlyVa 153
   |||||
378 GCACATCCGCGACGTACGATTTGGGGGACGCTACATCAGCCAGCAGCTTC 427
   |||||
153 HisPheMetSerLeuGlnIleGlyArgValMetGly.....LeuC 168
   |||||
428 CTTTCCCGTCGACCGCCATGTACAAACCGCGAAATCAAAAGCATAGAC 477
   |||||
168 yGlnPheMetMetAlaThrTyrArgProHisAsnGlnLeuMetGlu 184
   |||||
478 AAAATCATGACGGCGGAGGTTGCG...GGCAAGGAAACCGGCGCC 524
   |||||
185 TrpValGlnThrArgLysArgMetArgSerAsnLysAlaMetIleGlyAr 201
   |||||
525 TACGACGATACAGGGGTCAAAACATCATCAAAACCGCTGCTGGCGG 574
   |||||
201 gAsnAsnLeuArgLys.....IleValIleAlaLeuLysLysGly 215
   |||||
575 AAGCAACCATGCTCTGCGCCGACACGCTCCCTCCCTCAAGAAAGCGGG 624
   |||||
215 IuAlaValIleTrpPheAlaProAsp.....GlnAspTyrGly 226
   |||||
625 GAA.....GGCGTATGGTGATTTCTTCTGGC...AAACCTGCATATAC 665
   |||||
227 ArgLysGlySerSerPheAlaProPhePheAlaValAlaGlnAsnValAla 243
   |||||
666 CATGACCGCTGCGGCAAAATTTGCGACACGCTCAAAAGCGTGAAGAAACCTGT 715
   |||||
243 rThrAsnGlyThrTyrValLeuSerArgLeuSerGlyAlaAlaIleMetLeu 260

```

```

716 TTTTCTGCTGGCAACGCGCTGCGGACGACAGATTGCTGATTCACATC 765
   |||||
260 hrValThrMetValArgLysAlaAspTyrSerGlyTyrArgLeuPhe 276
   |||||
766 CGCCCGCTCCAAAGGGGATTCAGCGC.....GACAAAGCCCATGA 806
   |||||
277 ThrPro.....GlnMetClnIleTyrProThrAspGlnAsnGlnAl 290
   |||||
807 TGCCCGCGCTGTTCAACCGCAATGCCGATATGATGATGCGCGTTTCCGA 856
   |||||
290 alaAlaLysMetAsnLysIleIleGluLysGluIleMetArgAlaProG 307
   |||||
857 CGCAGTATCTGTTATGTACAAACCGCTACAAA 888
   |||||
307 IuGlnTyrLeuThrIleHisArgArgPheLys 317

```

seq\_name: p1r2:AD0637

seq\_documentation\_block:

lipid A biosynthesis [auroy] acyltransferase (EC 2.3.1.-) (heat shock protein) [Impor  
 C:Species: salmonella enterica subsp. enterica serovar Typh  
 A:Note: this species has also been called salmonella typh  
 C:date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
 C:accession: AD0637  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
 , S.; Mouton, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
 A:title: Complete genome sequence of a multiple drug resistant salmonella se  
 A:Reference number: AB0502; PMID:11677608  
 A:Accession: AD0637  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-306 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD08279.1; PID:916502326; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STY1192  
 C:Keywords: acyltransferase

alignment\_scores:  
 Quality: 185.00 Length: 289  
 Ratio: 1.194 Gaps: 8  
 Percent Similarity: 53.633 Percent Identity: 24.221

alignment\_block:

US-09-303-518D-569 x AD0637 ..

Align seg 1/1 to: AD0637 from: 1 to: 306

```

88 CTGGCGCTTCTGTCGACACGCTGGGAAACGGCTGCGACATCTGGC 137
   |||||
32 LeuProTyrProValIleTyrLysLeuGlyCysAlaLeuGlyHisLeuAl 48
   |||||
138 GTTTTACCTTTTAAAG..... 153
   |||||
48 aaArgArgValMetLysArgArgAlaLysIleAlaLysArgAsnLeuGluL 65
   |||||
154 .....GAAGACCGCGCGCGCATCGTCGCC 177
   |||||
65 eucysPheProGlnMetSerAlaGlnIleArgHisThrMetValValLys 81
   |||||
178 AATATGCGCTGACGAGCGATGAATCCGACCCCAAAACGGTCAAGCCGT 227
   |||||
82 AsnPheGlnSerValGlyMet..... 88
   |||||
228 TTTTCCGGAACGCAAAAGCGGCTTTGGAATTCGCCCGGCTTTTCA 277
   |||||
89 .....GlyValMetClnThrGlyMetAlaIlePheT 99
   |||||
278 GAAACCGGAGACATAGTAACAAATGTTCAAAAGCGGTACACGGCTGGAA 327

```



849 TTTCCGACGACATCTGTTATG 873  
 295 oalaprrogluIntrValtriple 303  
 seq\_name: p1r2:E71569

seq\_documentation\_block:  
 Probable acyltransferase - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C:Species: Chlamydia trachomatis  
 C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
 C:Accession: E71569  
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
 A:Reference number: A71570; MUID:99000809  
 A:Accession: E71569  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-455 <ARN>  
 A:Cross-references: GB:AE001275; GB:AE001273; NID:g3328388; PIDN:AAC67600.1; PID:g332839  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: htrB

alignment\_scores:  
 Quality: 172.50 Length: 342  
 Ratio: 1.033 Gaps: 12  
 Percent Similarity: 48.830 Percent Identity: 23.977

alignment\_block:

US-09-303-518d-569 x E71569 ..

Align seg 1/1 to: E71569 from: 1 to: 455

```

22 CGTTTCCCTTGGCAACGGCCATGCACATCCCTGTGACCCCTG.. 69
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
2  LeuphelsArgLeuArgThrGlyLeuValAspHisLeuVal 18
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
70 .....CTCAAAATGCCTCTCCCTGCGCGCTT 97
   ||||| ||||| ||||| ||||| ||||| |||||
18  lTyrGlyLeuGlyLeuGlyValLeuThrIleLeuArgLeuProAla 35
   ||||| ||||| ||||| ||||| ||||| |||||
98 CCTGCTGCACACGCTGGAAACGGCTCGACATCTGGCTTTTACCTT 147
   ||||| ||||| ||||| ||||| ||||| |||||
35  erSerLeuArgLeuPheSerGlyLeuGlyThrAlaLeuPheTyrphe 51
   ||||| ||||| ||||| ||||| ||||| |||||
148 TTAAGGAAGACCGCGCGCATGCTGCCAATATG..... 183
   ||||| ||||| ||||| ||||| ||||| |||||
52  IleSerAspValArgLysThrAlaLeuThrAsnLeuAlaLeuAlaPhePr 68
   ||||| ||||| ||||| ||||| ||||| |||||
184 .....CGTCAGGCA..... 192
   ||||| ||||| ||||| ||||| ||||| |||||
68  oGluSerPheAlaGluArgTyrGlnIleAlaArgGlnSerValGln 85
   ||||| ||||| ||||| ||||| ||||| |||||
193 .....GGCATGAATCCCGACCCCAAAACGGCTGAA 222
   ||||| ||||| ||||| ||||| ||||| |||||
85  lmetIleIleThrPheValGluLeuAlaThrValAspLysPheAlaLys 101
   ||||| ||||| ||||| ||||| ||||| |||||
223 GCGCTTTTGGGAACGGCAAAAGCGGTTTGAACCTGGCCCC...GC 269
   ||||| ||||| ||||| ||||| ||||| |||||
102  HisIleAspGluMetIleAlaIleAlaThrSerGlnAspAlaProGlu 118
   ||||| ||||| ||||| ||||| ||||| |||||
270 GTTTTTCGAACCGGAGACATAGAAACAATGTTCAACGGCTACAG 319
   ||||| ||||| ||||| ||||| ||||| |||||
118  yPhePhe.....ProGluGluValSerGlnGlnGluLeu..... 130
   ||||| ||||| ||||| ||||| ||||| |||||
320 GCTGGAAACATGTGACGAGCTTTGACAAACGAAAGGCTGCTATTC 369
   ||||| ||||| ||||| ||||| ||||| |||||
131 .....AspHisPhePheSerArgLeuAspArgGlnGluValAlaIleLeu 145
   ||||| ||||| ||||| ||||| ||||| |||||
370 ATCAGCGCGCACATCGGACGTACGATTTGGGGGAGCGTACATCAGCA 419
   ||||| ||||| ||||| ||||| ||||| |||||

```

```

146  PheCysGlyHisGlnAlaAsnTyrGluLeuProPheLeuTyrIleThr 162
420  GCAGCTTCCGTTCCCGCTGACCCCATGTACAAACCGCGAAATCAAG 469
   ||||| ||||| ||||| ||||| ||||| |||||
162  sArg.....TyrProGlyLeuAlaPheAlaLysProValLysAsnArg 177
   ||||| ||||| ||||| ||||| ||||| |||||
470  CGATGACAAATATCATGCAGCGCGGAGGTTCCCGCAAGAAACACC 519
   ||||| ||||| ||||| ||||| ||||| |||||
177  rGluAsnGlnLysIleIleSerLeuArgGluSerPheGlnGlyLys 193
   ||||| ||||| ||||| ||||| ||||| |||||
520  GCGCTTACACATACAAAGGGTCAACAAATCATCAAGCCCGCGCTC 569
   ||||| ||||| ||||| ||||| ||||| |||||
194  ValProGlu.....AsnAlaIleAsnGlnAlaLeuArgAlaHisArg 209
   ||||| ||||| ||||| ||||| ||||| |||||
570  GGGCGAAGCAACATCGCTCTGCGGACACGTCCTCCCTCAAGAG 619
   ||||| ||||| ||||| ||||| ||||| |||||
209  gGlyGluValValGlyLeuValGlyAspGlnValLeuLeuSerSerGlu 226
   ||||| ||||| ||||| ||||| ||||| |||||
620  GCGGGGAAGGGTATGGTGATTTCTTCGCAACCTGCTATACATG 669
   ||||| ||||| ||||| ||||| ||||| |||||
226  yrSer.....TyrProLeuPheGlySerGlnAlaPheThr 238
   ||||| ||||| ||||| ||||| ||||| |||||
670  AGCTGGCGCAAAATTCGACACGTCAAAGCGTGAACCCGTGTTT 719
   ||||| ||||| ||||| ||||| ||||| |||||
239  ThrSerProAlaLeuAlaTyrLysThrLysLysThrValIleAlaVal 255
   ||||| ||||| ||||| ||||| ||||| |||||
720  CTGCTCGAAGCGCTGCTGGCGGACAGGTTTGATTTGCATCGCGC 769
   ||||| ||||| ||||| ||||| ||||| |||||
255  lAlaIleTyrArgLysProAsnGlyAsnTyrLeuVal..... 268
   ||||| ||||| ||||| ||||| ||||| |||||
770  CCGTCAAGGGGAATGACGGGACAAAGCCCATGATGCCCGCGTTC 819
   ||||| ||||| ||||| ||||| ||||| |||||
269 .....ProSerLysAlaPhe 273
   ||||| ||||| ||||| ||||| ||||| |||||
820  AACCGCAATCCGCAATATGATGATACG..... 846
   ||||| ||||| ||||| ||||| ||||| |||||
274  HisAlaAsnThrGluLeuSerIleArgGluSerThrGluGlnLeuMet 290
   ||||| ||||| ||||| ||||| ||||| |||||
847 .....CGTTT.....CCGAGCGCAGT 862
   ||||| ||||| ||||| ||||| ||||| |||||
290  pArgLeuMetArgPheLeuGluLysGlyIleThrCysLysProGluGln 307
   ||||| ||||| ||||| ||||| ||||| |||||
863  ATCGTTTATGTACAAACCGCTACAA 888
   ||||| ||||| ||||| ||||| ||||| |||||
307  rPheuThrPheHisLysArgTyrLys 315
   ||||| ||||| ||||| ||||| ||||| |||||

```

seq\_name: p1r2:E71631

seq\_documentation\_block:

l1pid A biosynthesis lauroyl acyltransferase (htrB) RP718 - Rickettsia prowazekii  
 C:Species: Rickettsia prowazekii  
 C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
 C:Accession: E71631  
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark Nature 396, 133-140, 1998  
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
 A:Reference number: A71630; MUID:99039499  
 A:Accession: E71631  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-290 <AND>  
 A:Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA15149.1; PID:e134  
 A:Experimental source: strain Madrid E  
 C:Genetics:  
 A:Gene: htrB; RP718

alignment\_scores:  
 Quality: 169.50 Length: 206  
 Ratio: 1.274 Gaps: 9  
 Percent Similarity: 64.563 Percent Identity: 22.816

alignment\_block:



US-09-303-518D-569 x E71631 ..

Align seg 1/1 to: E71631 from: 1 to: 290

```

289 GACATAGAACAAATGTTCAAGCGGTACAGCGGTGGAGACATGTGCACA 338
95 GIUENGLUSERARGILEGLIUIELLE...GLYLSGLUSNILEUYSLY 110
339 GCGTTTGACAAACAGAGGGCTGTATTCATCAGCCGACATCGGCA 388
110 SLEUGLNASP.....GlnProPheLeuPheSerGlyHisPheAla 125
389 GCTAGATTTGGGGGAGCGGTACATCAGCCGACGCTTCCGCTCCGCTG 438
125 snTPASPILeserLeuYsValLeuYsLysTYrPro..LysVal 140
439 ACCGCGATGATCAAAACCGCCGAAATCAACGATAGCAAAATCATGCA 488
141 AlAValIleTYrArgLYsAlaAsnAsnProTYrValAsnLYsLeuValAs 157
489 GCGCGGAGGGTTCGCGGCAAGAAAACCGCGCTACCGCATACAG 538
157 nGLuserArgAlaGLysAspLYsLeuArgLeuIleProLYsGLysSerLUG 174
539 GCGTCAACAATCATCAAAAGCCCTGCTGCGGCGAAGCAACCATGCTC 588
174 LYIleArgSerLeuValArgAlaIleLYsGLUGLYserIleValMet 190
589 CTGCCCCAGCACGTCCTCCCTCAGAGAGCGGGAGAGCGTATGGT 638
191 LeuValAsp.....GlnLYsMetAsnAspLYsIleGLuVal 202
639 GGATTTCTCGGCAACCTGCTATACATGACGCTGCGGCAAAATG 688
202 LProPheLeuGLyArgProSerMetThrAlaSerIleAlaLYsIleA 219
689 CACACGTCAAAAGCGTGAAGCCCTGTTTCTGCTGCGAAGCGCTG... 735
219 La.....LeuGLnTYrLYsTYrProIleIleProCYsGLnIleLeuArg 233
736 CCTGCGGCAAGATGTTTGCATTTGCACATCGCGCCC...GTCCAAAGGGA 782
234 ThrLYsGLysSerTYrPheLYsValIleValHisProGLnLeuLYsPheGl 250
783 ATTGAACGCGCAAAAGCCCATGATGCC.....GCCGTGTTCAACCGCA 826
250 uGLnThrGLysAsnLYsValAspCYsTYrAsnIleMetPheAsnIleA 267
827 AT.....GCCGAATATTGGATACGCCGTTTCCGACGCGATATCTGTT 870
267 snGLnIleLeuGLysLysIleTYrValLYsGLnAsnProAlaGLnIleProPheTYrP 283
871 ATGTACAACCGCTACAAA 888
284 PheHisAsnArgTYrLYs 289

```

seg\_name: plir2:F81720

seg\_documentation\_block:

```

1) plid A biosynthesis lauroyl acyltransferase, probable TC0278 [Imported] - Chlamydia mu
C:Species: Chlamydia muridarum, Chlamydia trachomatis Morn
C:Date: 31-Mar-2000 #sequence revision 31-Mar-2000 #text change 11-May-2000
C:Accession: F81720
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis Morn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255
A:Accession: F81720
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-453 <TEF>
A:Cross-references: GB:AE002295; GB:AE002160; NID:g7190314; PID:AAF39146.1; PID:g719031

```

A:Experimental source: strain Nigg (Morn)  
C:Genetics:  
A:Gene: TC0278

alignment\_scores:  
Quality: 164.50 Length: 336  
Ratio: 0.962 Gaps: 14  
Percent Similarity: 50.893 Percent Identity: 24.702

alignment\_block:  
US-09-303-518D-569 x F81720 ..

Align seg 1/1 to: F81720 from: 1 to: 453

```

22 CTGTTTCCCCCTTGGAGACCGGCATGCATCTGTTG..... 60
2 LeuPheLYsMetLeuArgSerGLyLYsValLeuValAspHisLeuVal 18
61 .....ACCGCCTGCTCAATGCTCTCCCTGCTGCGGCTT 97
18 lYrGLyLeuGLyLeuGLyIleLeuThrIleLeuArgLeuLeuProArgS 35
98 CCTGTCTGCACACGCTGGGAAACCGGCTGCGACATCTGCGCTTTACCT 147
35 erSerLeuGLnLeuPheGLyLYsGLyLeuGLyThrThrIlePheLYrAl 51
148 TTAAGAGAACCGCGCGCGCATGTCGCCAATATGCTCAGCAGGCAAT 197
52 lIleSerAspPheArgLYsThrAlaLeuThrAsnLeuAlaLeuAla.... 66
198 GAATCCCGACCCC.....AAACGG 217
67 .PheProGLnLYsSerPheThrGLnArgTYrGLnIleAlaLeuLYsSerV 83
218 TCAAGCGCGT.....TTGCGGAAACGCAAA..... 246
83 alGLnGLnValIleIleThrPheIleGLnLeuAlaThrValAspLYsPhe 99
247 .....GGCGTTTGGAACTTCCCC 266
100 AlAlYsHisIleAspGLnIleIleThrIleAlaSerSerGLnAspAlaBr 116
267 C...GCGTTTTCAGAAACCGGACAGACATAGAAACATGTTCAACGCG 313
116 oGLnGLyPhePhe.....ProGLnGLnValSerSerGLnGLnLeu. 130
314 TACACGCGCTGGAAACATGTGCACAGCGCTTTGGACAACAGAGGGCTG 363
131 .....AsnAsnPhePheSerArgLeuAsnProGLnGLnGLnLYsAla 143
364 CTATTTCACAGCGCGCACATCGGCAGCTACGATTTGGCGAGCGTACAT 413
144 lIleLeuPheCYsGLyHisGLnAlaAsnTYrGLnLeuProPheLeuTYr 160
414 CAGCGACGACCTTCCGCTCCGTCAGCCGCTATCAAAACCGCGAAA 463
160 eThrLYsArg.....TYrProGLyLeuAlaPheAlaLYsProValLYsA 175
464 TCAAGCGATAGACAAATCATGCGAGCGGAGGCTTCGCGCAAGGA 513
175 snProArgLeuAsnArgLYsIleIleSerLeuArgLysPheGLnGLy 191
514 AAAACCGCGCTTACCATCATCAAGAGGGTCAACCAATCATCAAGCGCT 563
192 LYsIleValProProGLn...AsnAlaIleAsnGLnAlaLeuArgAla 207
564 GCGTTGCGGGAACCAACCATGCTGCTGCGGACGACGCTCCCTCCCTC 613
207 uHisLYsGLyGLnValValGLyIleValGLyAspGLnValLeuLeuSerS 224
614 AAGAGCGGGGGAAGGCGTATGGGTGATTCTTGGCAAAACCTGCTAT 663

```

```

224 erGlnTyrSer.....TyrProLeuPheGlySerGlnAlaPhe 236
664 ACCATGACGCTGGCGCAAAATGTCACACGCAAGCGGTGAACACCT 713
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
237 ThrThrThrSerProAlaLeuLeuAlaTyrLysThrArgLysProValI 253
714 GTTTTTCCTGCTGGACGCTGCTGGCGGACA..... 747
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
253 eAlaValAlaIleTyrArgGlnProAsnGlyAsnTyrLeuValProS 270
748 .....GGTTTCGATTTTCACATCCGCCCTCCAAAGG 780
270 erLysAlaPheTyrAlaAsnThrGlnLeuSerIleArgLysSerThrGlu 286
781 GAATTGAACGGCGCAAGCCCATGATGCGCGCTGTTCACCGCAATGC 830
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
287 GlnLeuMet...AspLys.....LeuMetArgPheLe 296
831 CGAATATTGGATAGCGCGCTTTTCGACGCAATCTGTTATGTAACAC 880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
296 uGlnLysGlyIleAlaCysLysProGlnGlnTrpLeuTrpLeuHisLysA 313
881 GGTACAAA 888
313 rqrTrpLys 315

```

seq\_name: p1r2:AG0251

```

seq_documentation_block:
  lipid A biosynthesis acyltransferase (EC 2.3.1.-) [Imported] - Yersinia pestis (strain O
  C:Species: Yersinia pestis
  C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
  C:Accession: AG0251
  R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
  deno-Terraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
  11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
  Nature 413, 523-527, 2001
  A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
  A:Reference number: AB0001; MID:21470413; PMID:11586360
  A:Accession: AG0251
  A:Status: preliminary
  A:Molecule type: DNA
  A:Residues: 1-320 <KUR>
  A:Cross-references: GB:AL590842; PIDN:CAC90875.1; PID:g15980074; GSPDB:GN00175
  C:Genetics:
  A:Gene: msbB
  C:Keywords: acyltransferase

```

```

alignment_scores:
  Quality: 158.50      Length: 288
  Ratio: 1.023         Gaps: 12
  Percent Similarity: 53.819      Percent Identity: 23.958

```

alignment\_block:  
US-09-303-518d-569 x AG0251 ..

Align seg 1/1 to: AG0251 from: 1 to: 320

```

70 CTCGAATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCAACGGGGA 119
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
34 IleAlaIleAlaLeuAlaTyrIleProProLysPheArgAspProLeuAla 50
120 CCGGCTGGACATCTGGCGTTTACCTTTAAAGAAAGACCGCGCGCA 169
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 aglyIleLeuArgLeuAlaGlyLysPheAlaLysSerAlaArgArgArg 67.
170 TCGTCGCAAT.....ATGCGTCAGGACGACGATGATCCCGAC 207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 IaaArgIleAsnLeuLeuTyrCysMetProGluLeuProGluSerGluArg 83
208 CCCAAACGGTCAAAAGCGTTTTCGCAAAAGCGCAAAAGCGGTGGA 257
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

84 GlnHisIleIleAspGlnMetPheAlaThrAlaIleGlnProLeuMetMe 100
258 ACTGGCCCCCGCGTTTTCAGAAAACCGGAGACATAGAAACAATGTTCA 307
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
100 tMetAlaIleLeuLysPheArgAspProLysLysValLeuThr..... 114
308 AAGCGTACAGCGCTGGACATGTGACAGCGCTTGGAC..... 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
115 ..ArgValHis...Trp...HisGlyGlnIleLeuAspGluLeuGln 128
349 ..AAACAGAAAGGCTGCTATTCATCAGCGCGCACATCGGACGTACGA 395
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129 GlnGlnIleArgAsnValIleLeuLeuValProHisAlaTrpSerIleAs 145
396 TTTGGGGGAGCTTACATCAGCAGCAGCTTCCGTCGCGTACCGCA 445
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 rIleProAlaMetLeuLeuAlaGlnGln...GlyLysProValAlaGlyM 161
446 TGTCAGAAACCCCGCAAAATCAGAACGATAGCAAAATCAGACGGCGGC 495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
161 eTrpHisHisGlnArgAsnProLeuValAspTyrLeuTrpAsnSerAla 177
496 AGGCTTCGCGCAAGGAAAAACCGCGCTACACGATACAGGCGTCA 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
178 ArgLeuHisPheGlyGlyArgIleHisAlaArgGln...SerGlyIle 193
546 ACAATATCATCAAGCCCTGCTGGCGGAGAACCATGCTGCTGCCG 595
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
193 sProPheIleSerSerValArgGlnGlyPheTyrGlyTyrTyrLeuPro 210
596 ACCAGCTCCCTCCCTCAGAAAGGCGGAGCGGATGCGGTGCTTC 645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 spGlnAspTyrGlyProGln.....GlnSerGluPheValAspPhe 223
646 TTGGCAACCTGCTATACCATGACGCTGCGCAAAATGTCACACGT 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
224 PheAlaThrTyrLysAlaThrLeuProAlaIleGlyArgLeuMetLysA 240
696 C.....AAGCGGTGAACCTGTT.....TTTGCTGCGG 727
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
240 LysAlaGlnAlaIleValProMetPheProValTyrAsnTyrArgGln 257
728 AAGCGCTGCTGCGGACGAGGATTTGATTTGACATCGCGCCGTCGAA 777
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
257 tsArgLeu.....AspIleTyrIleArgPropromet 267
778 GGGGAATG...AAGCGGACAAAGCCCATGATGCGCGCTGTCAACG 824
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
268 AspAspLeuAlaAspAlaAspAlaTyrIleAlaArgArgMetAsnG 284
825 CAATGGCGAATATTGATACCGCGCTTTCGACGCAATCTGTTATGT 874
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
284 uGlnValGlnLeuLeuValLysProAsnProGlnTyrThrTrpIleL 301
875 ACAACCGCTACAA 888
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 euLysLeuLeuLys 305

```

seq\_name: p1r2:AE3391

```

seq_documentation_block:
  lipid A biosynthesis lauroyl acyltransferase (EC 2.3.1.-) [Imported] - Brucella melit
  C:Species: Brucella melitensis
  C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
  C:Accession: AE3391
  R:DeLVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
  .; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
  Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
  A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit
  A:Reference number: AD3252; PMID:11756688
  A:Accession: AE3391
  A:Status: preliminary
  A:Molecule type: DNA

```

A;Residues: 1-307 <NR>  
 A;Cross-references: GB:AE008917; PIDN:AA152296.1; PID:g17983087; GSPDB:GN00190  
 A;Experimental source: strain 16M  
 C;Genetics:  
 A;Gene: BME1115  
 A;Map position: 1  
 C;Keywords: acyltransferase

alignment\_scores:  
 Quality: 152.50 Length: 330  
 Ratio: 0.902 Gaps: 13  
 Percent Similarity: 51.212 Percent Identity: 23.636

alignment\_block:  
 US-09-303-518D-569 x AE3391 ..

Align seg 1/1 to: AE3391 from: 1 to: 307

```

1 ATGTTGCTTACATTCAGGCTGTTCCCTTGGCAGCCGATGCA 50
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MetPheLysLeuLysLeuLeuLeuPheArgTrpSerArgLysLeuLysGcl 17
51 CATC.....CTGTTGACGCGCTGCTCAATGCC 79
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 nPheasnTrpLeuTrpAlaGlnAlaValPheValLeuLeuGlyPheL 34
80 TCTCCCTGCTGCGCTTCTGCTGTCACAGCGTGGAAACCGCTCGA 129
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 euArgLeuPheProAlaLysAlaAlaLysSerPheSerAlaArgValAla 50
130 CATCTGGGGTTTACCTTTTAAAGAGACGCGCGCATCGTGGCCAA 179
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 ArgLeuLleGlyProLeuThrProArgHisLysVal...AlaThrAspAs 66
180 TATGGGTGACGAGCATCATGATCCGACCCCAAAAGCGTCAAGCGCTT 228
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 neuArgLysAla....TyrProGluLysSerGluAlaGluLeuValL 81
229 .....TTTGGGGAACG 240
81 ysLleAlaArgGluMetTrpAspSerMetAlaArgLeuPheAlaLutyr 97
241 GCAAAAGCGGTTTGAACTGCCCCGCTTTTCAGAAAAACCGAAGA 290
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
98 .....IlePheLeuAspAlaValPheAspPheAspProTyrAl 110
291 CATAGAAACATGTTCAAGCGGTACAGCGCTGGAAACATGTGCAGCAGG 340
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
110 aValLysProGlyLeuValGluValGluLysLeuProLlePheGluGlyL 127
341 CTTTGGACAAACAGCAAGCGGTATTCATCAGCGCGCATCGGAGCAGC 390
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
127 euArgAspLysLysProHisLlePhePheThrAlaHisThrLysLysN 143
391 TAGGATTTGGCGGAGCGTACATCAGCCAGCAGCTTCG..... 429
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
144 PheGluLeu.....LeuProLleCysAlaAl 152
430 .....TTCCCGCTGACCGCATGTACAAACCGCGGAAATCAAG 469
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
152 aThrPheGlyLeuAsnValThrAlaLeuPheArgProProAsnAsnProT 169
470 CGATA...GACAAATCATGACAGCGGCGGCGGCTCGCGCAAGAA 516
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 yTrlLeuAsnLysValLeuLysAlaArgArgThr...AsnMetGlyHis 184
517 ACCGCGCTACAGCATTACAGGGGTCAAAACATCATCAAGCCCTGCG 566
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
185 LeuValProSerLysAlaGlyAlaAlaTrpAlaLeuAlaGlyLleLeu 201
567 TTGCGGCGAAGCAACCATGCTGCTGCGCCACACACGTCCTCCCTCAG 616
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

201 ysArgGlyLysAsnValGlyMetLeuValAsp.....GlnL 213
617 AAGCGCGGAGAGCGCTATGGATTTCTTGGGCAACCTGCTATAC 666
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 ysPheSerArgGlyValProSerThrPhePheAsnArgProValLysThr 229
667 ATGACCGTGGCGGCAAAATTTGGCACAGCTCAAGCGGTGMAACCTGTT 716
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
230 AsnProLeuLeuAlaLysLeuAlaArgGlnTrpAspCysAspValLysPr 246
717 TTTCGCTGCGAAGCGCTGCGGAGACAAGTTTGCATTTGCACATC 765
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
246 oAlaArgCysLleArgLeuProGlyGlyArg...TyrArgLeuGluLeuT 262
766 .....CGCCCGCTCAAGAGGAGATTTGACGCGCAGC 795
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
262 yGluArgMetGluLeuProArgAspGluSerGlyGlnLle.....Asp 276
796 AAGCCCATGATGCGCGCTTTCACAGCGCAATGCCGAATATTGATACG 845
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
277 IleAsnAlaThrThrGlnLeuLeuAsnAspThrValGluGlnTrpValAr 293
846 CCGTTTCCGACGACATATCTGTTATGTACAAACGCTAC 885
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
293 gGluTyrProGlyGlnTrpMetTrpPheHisLysArgTrp 306

```

seq\_name: p1r2:T42129

seq\_documentation\_block:  
 probable acyltransferase (EC 2.3.1.-) - Escherichia coli plasmid p0157  
 C:Species: Escherichia coli  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Nov-2000  
 R:Buriland, V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.  
 Nucleic Acids Res. 26, 4196-4204, 1998  
 A:Title: The complete DNA sequence and analysis of the large virulence plasmid of Esc  
 A:Reference number: Z22068; MUID:96391744  
 A:Accession: T42129  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-343 <NR>  
 A:Cross-references: EMBL:AF074613; PIDN:ANC70097.1  
 A:Experimental source: Strain EBL933; serotype O157:H7  
 R:Makino, K.; Ishii, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubo  
 S.; Shinagawa, H.  
 DNA Res. 5, 1-9, 1998  
 A:Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemor  
 A:Reference number: Z14127; MUID:98290540  
 A:Accession: T00321  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 12-343 <MAK>  
 A:Cross-references: EMBL:AB011549; NID:94589740; PIDN:BA031840.1; PID:93337081  
 A:Experimental source: strain EHEC O157:H7, substrain RIMD 0509592  
 C:Genetics:  
 A:Genome: plasmid p0157  
 A:Note: L7029  
 C:Keywords: acyltransferase

alignment\_scores:  
 Quality: 151.00 Length: 286  
 Ratio: 0.993 Gaps: 9  
 Percent Similarity: 53.147 Percent Identity: 22.028

alignment\_block:  
 US-09-303-518D-569 x T42129 ..

Align seg 1/1 to: T42129 from: 1 to: 343

```

82 TCCCTGCTGCGCTTCTGCTGTCACAGCGTGGAAACCGCTCGACA 131
  |||||:|||||:|||||:|||||:|||||:|||||:|||||:
67 AlaMetValProProAlaLeuArgAspProLeuLeuLysLysLeuGlyMe 83

```

```

132 TCTGGCGCTTTACCTTTAAAGAGACCGCGCGCATGCTGGCCAAAT 181
133 ||||| ||||| ||||| ||||| ||||| ||||| |||||
83 tleuvalglargleuclylserhisalarginalaleuileuileu 100
182 TGGCT.....CAGCAGCATGATGATCCGACCCCAAAAGCGTC 219
100 enserleucyspneproglutylseraspysglulysgluansileval 116
220 AAGCCGCTTTTTCGGAAACGGCAAAAGCGCTTGGAACTTGCCCGCC 269
117 Aspalamephelealathralasmetalaivalleuemetalaile 133
270 GTTTTTCAGAAACCGGAACATACAAACATGTTCAAGCGGTACAG 319
133 ualaleuserglyproaspysileserhisargilearg...TTPasn 149
320 GCTGGAGACATGTCAGCAGCGCTTGGCAAAACAGCAGCGCTATTTC 369
149 lylleuglulilevalglulysmetalaileuansnleuylsle 165
370 ATCAGCCCGCATCGCAGCTACGATTG..... 399
166 leuvalprohisalatrprglyalasprieproalileuemetala 182
400 .GGCGAGCCTACATCAGCCAGCAGCTTCCGCTGACCGCATGT 448
182 aserglyarg.....lysmetalaileu 191
449 ACAAAACCGCGCAAAATCAAGCATAGACAAATCATGACGCGCGG 498
191 henhisanslnargasnprovalalasptryvalitprasnervall 207
499 GTTCCGCGCAAAAGCAAAACCGCGCTACACCATACAAAGGGTCA 548
208 arguargpnehllyllysleuhsilalargasn...Aspglylale 223
549 AATCATCAAGCGCTGCGTGGGGAAGCAACCATGCTGCTGCC 594
223 rthevalargservalarglingleltytrpelytyrtyrleup 240
595 ..GACCAGTCCCTCCCTCAAGAGCGGGAAGCGGTATGGGTGAT 642
240 lnasphlgltyproglupheserlu.....phealasp 251
643 TCTTCGCGCAAACTGCTATACATGACGCTGGCGCAAAATGGCACA 692
252 phehealathrtyrlysalathtleuuprovalileglyargleu 268
693 CGTCAAGCGGTGAAGACCTGTGTTCTGCTGGCAAGCGCTGGCG 742
268 gileserglyalargilelle.....Proleupheproval 281
743 GACAAGGT.....TTCGATTTCGACATCCGCCCGCTCAAGG 780
281 yraspglyllysthrhisleuthrllehisvalserproleu 297
781 GAATTGACGCGCAAGCCCATGATGCCCGCTTCAACCCGCAATC 830
298 lleruglulysseraspalalhisilealargclintleasnleu 314
831 CGAATATTGATACGCGCTTTCGACGACATATCTGTATTATGACA 880
314 lgluasrphlevalargprohisprogluglntlythrtripleu 331
881 GCTCAAA 888
331 euleuyls 333
seq_name: p1r2:C97836
seq documentation block:
hypothetical protein htrb [imported] - Rickettsia conorii (strain Malish 7)

```

```

C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: C97836
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; M01D:2142074; PMID:1157893
A:Accession: C97836
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-290 <KUR>
A:Cross-references: GB:AE006914; PIDN:AAL03629.1; PID:915620214; GSPDB:GN00173
A:Gene: htrb

alignment_scores:
  Quality: 149.00      Length: 201
  Ratio: 1.221        Gaps: 8
  Percent Similarity: 60.697      Percent Identity: 23.383

alignment_block:
US-09-303-518D-569 x C97836 ..

Align seg 1/1 to: C97836 from: 1 to: 290

310 GCGGTACACGCGCTGGAAACATGTGCAGCAGGCTTGGCAAAACGAGG 359
101 Alallelleglyllegluanslleys.....lysleuglul 113
360 G.....CTGCTATTCAATCAGCGCGCATCGCAGCTACGATGTTGGCG 403
113 yglprpheleueneurhegllyglhisphealasntrpasplesert 130
404 GACCTACATCAGCCAGCAGCTCCGCTCCCGCTGACCGCATGTACAA 453
130 lelysileleuhsilyspheytypro...lysvalalalaletyarg 145
454 CCGCGCAAAATCAAGCATAGACAAATCATGACGCGCGGAGGTTGC 503
146 lysalasnasnprotyrvalalasnlysleuvalasnleuvalargl 162
504 CGGCAAGCAAAACCGCGCTACACCATACAAAGGGTCAACAAATCA 553
162 yasplyleuargleuileleprolysglyproglugllyserarglaleu 179
554 TCAAGCGCTGCGTGGGGAAGCAACATGCTGCTGCCGCAACGCTC 603
179 alrpalalallelysleuergluserlilevalmetleuvalasp 193
604 CCCTCCCTCAAGAGCGGGAAGCGGTATGGGTGATTCTTGGCA 653
194 .....glnlysmetasnaspglyllegluvalprpheleu 207
654 ACCTGCTATACATGACGCTGGCGCAAAATGGCAACGTCAAAGGCG 703
207 sproalamethrhalaseralalealalysileala.....leuglnt 222
704 TGAAAACCGCTGTGCTGCGCAAGCGCTG.....CCTGCGCAAGGT 750
222 yrlslytyrproilleleleprocslelilleleargthrlysglyser 238
751 TTGCAATTGACATCGCGCC...GTCAAGGGGAATTGACGGGACAA 797
239 Pheylsvalillevalhisprogluileuylsphegluglntthglyasp 255
798 AGCCCATGATGCC.....GCCGTTCACACCGCAATGCCGAAT 835
255 nylsalaaspyslyrasnilemetleuanslleasnleuileu 272
836 ATTGATACGCGCTTTCGACGACATATCTGTATTATGACAACGCTAC 885
272 lntprvalylsleuansnproalaglnttrpphetrphhisasnarg 288

```

886 AAA 888  
111  
289 Lys 289

seq\_name: p1r2:D64127

seq\_documentation\_block:

lipid A biosynthesis lauroyl acyltransferase (EC 2.3.1.-) - Haemophilus influenzae (strain N) alternate names: htrb protein  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 08-Oct-1999  
C:Accession: D64127  
R:Reischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gockyne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: D64127  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-315 <TIGR>  
A:Cross-references: GB:U032828; GB:LA2023; NID:g1574362; PIDN:AAC23173.1; PID:g1574368; T C:Genetics:  
A:Gene: htrb  
C:Function: GTG  
A:Pathway: lipopolysaccharide biosynthesis  
A:Keyword: acyltransferase; lipopolysaccharide biosynthesis; transmembrane protein  
F:132-148/Domain: transmembrane #status predicted <TIGR>

alignment\_scores:

Quality:	Ratio:	Length:
148.50	0.946	287
Percent Similarity:	54.704	Gaps: 11
Percent Identity:		25.087

alignment\_block:

US-09-303-518D-569 x D64127 ..

Align seg 1/1 to: D64127 from: 1 to: 315

```

64 GCCCTGCTCAATGCTCTCCCTGCTCCCTGCTCTGTCACAGGCT 113
   ||||| : : : : ||||| : : : : :
30 AlaIleTPArGSeRIleLeuCySLeuPrOtrProIleuArGhISII 46
   ||||| : : : : ||||| : : : : :
114 GGGAACCGGCTCGGACATCTGGCGTTTACCTTTAAAGAGACCGCG 163
   ||||| : : : : ||||| : : : : :
46 eGIYhIScIYrheGIYrPrleuPheSerhISleuLYsValGIyGluArG 63
   ||||| : : : : ||||| : : : : :
164 CGCGATGCTGCGCAATATGCGTCAAGGACGATCAATCCGAC...CCG 210
   ||||| : : : : ||||| : : : : :
63 rGAlAlAlAllelAlArGArGySLeuGIuLeuCySrPhePrOArPhePrO 79
   ||||| : : : : ||||| : : : : :
211 AAAAGCGTCAAGCGCTTTTGGCGAA.....ACGGCAAA 245
   ||||| : : : : ||||| : : : : :
80 GIuAnGluArGyGluValIleLeuGIuAnSLeuArGSeArValGIyMe 96
   ||||| : : : : ||||| : : : : :
246 AGCGGTTTGAACCTTGCCCGCTTTTTCAGAAACCGGACAGCATAG 295
   ||||| : : : : ||||| : : : : :
96 ValAlleIleGIuYrGIyMeAlArTrPheTrPheSerArPSeArGIleI 113
   ||||| : : : : ||||| : : : : :
296 AAACAATGTTCAAGCGGTACAGCGGTGGAGACATGTGACAGACGCTTGG 345
   ||||| : : : : ||||| : : : : :
113 yAlYrTrPSeArYs...ValGIuGIuLeuNIstYrLeuLYsGIu..... 126
   ||||| : : : : ||||| : : : : :
346 GAACAACACGAGGCTGCTATTCACAGCGCGACATCGGACGTAGAG 395
   ||||| : : : : ||||| : : : : :
127 ASngInLYsArPDIYleValLeuValGIyValhISrPheLeuThrLeuGI 143
   ||||| : : : : ||||| : : : : :
396 TTTGGCGGACGCTACATGACGACGAGCTTCCGTTCCGCTGACCGGCA 445

```

```

143 uLeuCYlAlArGIleGIY.....LeuNIstYrPrOGLYleGIY 158
   ||||| : : : : ||||| : : : : :
446 TGTACAAACCGCGCAATTCAGAGTACAGCAAAATTCATCGACGCGG 495
   ||||| : : : : ||||| : : : : :
158 AlYrArGPrOArSArPArSArPrOleuArGArPTrPrleuGIYrheGIY 174
   ||||| : : : : ||||| : : : : :
496 AGCGTTCG.....GGCAAGGAAACCGCGCTTACGACATCAAGAGGCT 542
   ||||| : : : : ||||| : : : : :
175 ArGIeAlArGSeArSArLYsArMeLeuArPArGySArPLeuArGIY.. 190
   ||||| : : : : ||||| : : : : :
543 CAACAATTCATCAAGCGCTGCTCGGCGGAAACGACATCGCTCTGC 592
   ||||| : : : : ||||| : : : : :
191 .....MeIleYsAlAlArGIhISGIuGIuThrIleTrPrYrAlAr 205
   ||||| : : : : ||||| : : : : :
593 CGACACGCTCCCTCCCTCAAGAGCGGG.....GAAGCGGTATGG 636
   ||||| : : : : ||||| : : : : :
205 rOArPArS.....ASrYrGIYArGySArSAlAlArPhe 216
   ||||| : : : : ||||| : : : : :
637 GTGATTTCTTGGCAACCTGCTATGACATGACGCTGGCGGCA.... 681
   ||||| : : : : ||||| : : : : :
217 ValPrOArPheAlAlArPrOArPrhTrGySrTrhTrhGIYSeArYrTy 233
   ||||| : : : : ||||| : : : : :
682 .AAATTCGACACGTCAAGCGGTGAACCTGTTTCTGCTCGGAC 730
   ||||| : : : : ||||| : : : : :
233 rLeuLeuYsSerGIeArSArYrValIlePrOArPheAlAr 250
   ||||| : : : : ||||| : : : : :
731 GCCTGCTCGGACAGGTTTCATTCGACATC...CGCGCGTCCAA 777
   ||||| : : : : ||||| : : : : :
250 rGASArLYsArPDIYrGIYrValSerIleSerAlArPrOValArP 266
   ||||| : : : : ||||| : : : : :
778 GGGAATTCAGCGGACCAACCCCATGATGCCCGCTGTTCAACGGCA 827
   ||||| : : : : ||||| : : : : :
267 rPheThrASrPLeuGIuArPArLArAlAlArGArMeArSngInII 283
   ||||| : : : : ||||| : : : : :
828 TGCCGAATATTCGATACGCCGTTTCCGACGACATGCTGTATAGTCA 877
   ||||| : : : : ||||| : : : : :
283 eValGIuLYsGIuIleMeLYsArPLeuThrIleYrMeArPLeuNIhSA 300
   ||||| : : : : ||||| : : : : :
878 ACGGCTACAA 888
   ||||| : : : : ||||| : : : : :
300 rGArPrPheLYs 303

```

seq\_name: p1r2:T45134

seq\_documentation\_block:

hypothetical protein [imported] - Microbacterium ammoniaphilum (fragment)  
C:Species: Microbacterium ammoniaphilum  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
C:Accession: T45134  
R:Striebel, H.M.; Seeber, S.; Jarsch, M.; Kessler, C. Gene 172, 41-46, 1996  
A:Title: Cloning and characterization of the MamI restriction-modification system from  
A:Reference number: Z22923; MUID:96257250  
A:Accession: T45134  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-529 <STR>  
A:Cross-references: EMBL:X79027; NID:g984667; PIDN:CA55649.1; PID:g1679831  
A:Experimental source: ATCC 15354

alignment\_scores:

Quality:	Ratio:	Length:
147.50	1.209	293
Percent Similarity:	41.638	Gaps: 16
Percent Identity:		28.669

alignment\_block:

US-09-303-518D-569 x T45134 ..

Align seg 1/1 to: T45134 from: 1 to: 529

30 CCTTTGGAAACGCGCATGACATCTGTGACCGCGCTGCAATGCC 79

[illegible]

```

seq_documentation_block:
  lipid A biosynthesis lauroyl acyltransferase, probable CP0676 [imported] - Chlamydia
  C:Species: Chlamydia pneumoniae pneumoniae, Chlamydia pneumoniae
  C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
  C:Accession: E01551
  R:Read, T.D.; Brumham, R.C.; Shen, C.; Gill, S.R.; Heldelberg, J.F.; White, O.; Hickey,
  R.C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
  Nucleic Acids Res. 28, 1397-1406, 2000
  A>Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39
  A:Reference number: A81500; WUID:20150255
  A:Accession: E81551
  A:Status: preliminary
  A:Molecule type: DNA
  A:Residues: 1-462 <REA>
  A:Cross-references: GB:AE002225; GB:AE002161; NID:g7189583; PIDN:AAF38487.1; PID:g718
  A:Experimental source: strain AR39, HL cells
  C:Genetics:
  C:Gene: CP0676

```

[illegible]

alignment\_block;

Align seg 1/1 to: E81551 from: 1 to: 462

```

34 TTGGAAACGGCATGACATTCCTGGTGGACGGCCGAGTCATTCGGCTCTC 83
13 LeuGluAlaProLeuTyrTyrLeuValIserGlyIleIleAlaLeuCysAr 29
84 CCTCTCTCCGCTTCTCTGTCTGCACAGCTGGAAACCGAGCTCGACATC 133
29 gNtStHrProArgSerPheLeuThrGluLeuGlnTyrGlnTyrGlnIlePhe 46
134 TGGGGTTTACGTTTAAAGGAACCGCGCGCGAGATGCTGTCGCAATATG 183
46 euAlaPheTyrTleIleSerPheTyrGlnTyrStHrAlaLeuThrAsnLeu 62
184 .....CG 185
63 AlaLeuAlaPheProGlnTyrStHrPheAsnGlnAlaGlnTyrIleAlaAr 79
186 TCAGGCGGCGATAAATCCGCGACCCCAAAACGGTCAAACCGCTTTTGGCG 235
79 gGlnSerLeuGlnIleLeuIleIleThrLeuLeuGlnIleLeuAlaIleG 96
236 AAACGGCAAAAGCGGCTTGGAACTTGCCTCCCGCGGTTTTTC..... 276
96 IuGlnLeuValGlyAsnIleAspTyrLeuIleThrIleValThrSerSer 112
277 AGAAACCG.....GAAACAT 293
113 ArgAsnProTyrGlnTyrPheSerSerGlnGluValIleSerAsnGlnAspLe 129
294 AGAAACATGTTCAAAAGCGGTACAGCGCTGGGAACATGTGCAGCAGCTT 343
129 uGlnGluTyrThrPheTyr.....Asn 136
344 TGGACAACAACGAGGCGTGTATTTCATCAGCGCCACATCGGAGCTAC 393
136 euGlnIuLysGlnGlyLeuIleLeuPheCysGlnTyrGlnAlaIAsnTrp 152
394 GATTGGGCGGACGCTACATACAGCCAGCAGCTTCCGTTCCGCTGACCGC 443
153 GluLeuProPheLeuTyrTleThrLysAsn.....TyrProGluIleAl 167
444 CATGTACAACCGCGGAAATTCAAACGCGATGACAAATCATCATCAGCGG 493
167 apheAlaLysAlaIleLysAsnGlnAlaGlnSerLysLysIlePheAlaL 184

```





299 aSerGlnProGlnIntPrMetTrpIleHisLysArgTrpLys 313

seq\_name: p1r2:D97552

seq\_documentation\_block:

lipid A biosynthesis lauroyl acyltransferase (hrb) rp718 [Imported] - Agrobacterium tum  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: D97552  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: AF7359; PMID:11743194  
A:Accession: D97552  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-312 <KUR>  
A:Cross-references: GB:AE007869; PIDN:AAK87373.1; PID:915156679; GSPDB:GN00169  
C:Genetics:  
A:Gene: AGR\_C\_2930  
A:Map position: circular chromosome

alignment\_scores:  
Quality: 141.00 Length: 289  
Ratio: 1.007 Gaps: 14  
Percent Similarity: 48.443 Percent Identity: 23.183

alignment\_block:

US-09-303-518D-569 x D97552 ..

Align seg 1/1 to: D97552 from: 1 to: 312

```

190 GCAGCGCATGATCCGACCCCAAAAGCGTCAAAAGCGTTTTCGCGAAG 239
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 AlaglyIleArgAlaThrAspArgLeuAlaArgLeuIleGlyProLysTh 57
240 GCGCAAAAGCGCGTTG.....GAACCTGCCCCCGCTTTTCAGAA 280
    |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
57 rclYrghIstLysLeuMetLeuTyrAsnLeuAlaArgAlaPheProGluL 74
281 AACCGGACATAGAACAAATGTTCAAAAGCGGTACAGCGCTG..... 324
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
74 ystHrGluGlu...GluArgLeuAlaIleAlaMetAspSerTrpAlaasn 89
325 .....GAACATGTC..... 333
    |||:|||||:|||||:|||||:|||||:|||||:
90 MetGlyArgLeuAlaIleGluTyrValPheLeuAspArgLeuPheAspPh 106
333 ..... 333
106 eAsProGluLysAsnGluProGlyArgIleGluValGlnGlyThrSerT 123
334 .....CAGCAGCGTTTGGCAAAACAGCAAGCGGTCTTATCATACGCGG 378
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
123 hrPheLeuGluLeuAlaArgAspAsnProArgProPheIleValPheThrAla 139
379 CACATCGGCACCTAGATTTGGCGGACGCTACATCAGCCAGCAGCTTCC 428
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
140 HisSerGlyAsnPheGluLeu.....LeuPr 148
429 G.....TTCGCGTGCAGCCGCATCATGACAAACCCG 457
    |:|||||:|||||:|||||:|||||:|||||:|||||:
148 oValAlaGlySerAlaPheGlyLeuAspValThrValLeuPheArgProP 165
458 CGAAATATCAAGGATA...GACAAATCATCGAGCGGCGAGCGGTTCGC 504
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
165 roAsnAsnProTyrValAlaAspLysValPheAsnPheArgLysGluArg 181
505 GCGAAAGGAAACCGCGCTACAGCATACAGGCGTCAACAAATCAT 554
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
182 ...MetGlyAsnLeuValProSerHisAlaGlySerSerPheAlaLeuAl 197

```

```

555 CAAGCCCTGCGTTGGCGGAAGCAACCATGCTCTCCGCGACGACGTC 604
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
197 aArgGlnLeuGlnLysGlyGlyGlyValGlyValLeuValasp..... 211
605 CCTCCCTCAAGAGCGGGGAGGCGGTATGCGTATGCTTTCTTGGCAAA 654
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
212 .....GlnLysPheGlyLysGlyLeuThrTrpLysPhePheGlyLeu 225
655 CCTGCTTATACCATGACGCTGGCGGCAAAATTTGCACACGTCAAGGCGT 704
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
226 GluValArgThrAsnProLeuLeuAlaLysLeuValArg..... 238
705 GAAACCCCTGTTTCTGTC.....TGCAGACGCTGC 736
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
239 .....GlnPheAsnLysAspValTyrProAlaArgCysIleArgLeuP 253
737 CTGGCGGACAAAGGTTTCGATTTCGACATCCGCGCGTCGCAAGGGAATG 786
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
253 roAspAsnArg...TyrArgLeuGlnIleGluPro...LysValGluIle 267
787 .....AACGCGACAAAGCCCATGATGCCGCGCTGT 818
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
268 ProArgAspGlnLysGlyAsnValAspIleGlnAlaThrAlaGlnLeuLe 284
819 CAACCGCAATGCGGATATGATACGCGCTTTCCGACGACGATGTGT 868
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
284 uAsnAspLysValGlnIleSerTrpValArgGlnTyrProGlyGlnTrpLeu 301
869 TTATGTACAAACCGCTAC 885
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
301 rPtyrHisAspArgTrp 306

```

seq\_name: p1r2:AF2772

seq\_documentation\_block:

lipid A biosynthesis lauroyl acyltransferase [Imported] - Agrobacterium tumefaciens (  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AF2772  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McCl  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AF2772  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-312 <KUR>  
A:Cross-references: GB:AE006886; PIDN:AA142596.1; PID:917740022; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: msbB  
A:Map position: circular chromosome

alignment\_scores:  
Quality: 141.00 Length: 289  
Ratio: 1.007 Gaps: 14  
Percent Similarity: 48.443 Percent Identity: 23.183

alignment\_block:

US-09-303-518D-569 x AF2772 ..

Align seg 1/1 to: AF2772 from: 1 to: 312

```

190 GCAGCGCATGATCCGACCCCAAAAGCGTCAAAAGCGTTTTCGCGAAG 239
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
41 AlaglyIleArgAlaThrAspArgLeuAlaArgLeuIleGlyProLysTh 57

```

```

240 GCGAAAGCGGTTG.....GACTTGGCCCCGTTTTCAGAA 280
      |||
      |||
      |||
      |||
57 rGlyAghHisLysLeuMetLeuTyrAsnLeuAlaIgaAlaPheProGlu 74
      |||
      |||
      |||
      |||
281 AACCGAGACATAGAAACAAATGTTCAAAAGCGGTACAGCGCTGG..... 324
      |||
      |||
      |||
      |||
74 ysrHrGluGlu...GluArgLeuAlaIleAlaMetLaspserTrrAlaAsn 89
      |||
      |||
      |||
      |||
325 .....GACATGTG..... 333
      |||
      |||
      |||
      |||
90 MetGlyArgLeuAlaIgaIuTyrValPheLeuAspArgLeuPheAspH 106
      |||
      |||
      |||
      |||
333 ..... 333
106 eaSprProGluLysAsnGluProGlyArgIleGluValGluGlyThrSert 123
      |||
      |||
      |||
      |||
334 .....CAGCAGGCTTTGGACAAACAGAGAGGCTGCTATTTCATCAGCCG 378
      |||
      |||
      |||
      |||
123 hrPheLeuGluLeuAlaGAspAsnProArgProPheIleValPheThrAla 139
      |||
      |||
      |||
      |||
379 CACATCGCAGAGCTACGATTGGCGGAGCTACATCAGCAGCAGCTTCC 428
      |||
      |||
      |||
      |||
140 HisSerGlyAsnPheGluLeu.....LeuPr 148
      |||
      |||
      |||
      |||
429 G.....TTCCGCTGACCCGCTATGACAAACCC 457
      |||
      |||
      |||
      |||
148 oValaIgaLysrAlaPheGlyLeuAspValThrValLeuPheArgProp 165
      |||
      |||
      |||
      |||
458 CGAAATATCAAAAGCGTA...GACAAATCATGACGGGCGGAGGTTTCG 504
      |||
      |||
      |||
      |||
165 roAsnAsnProTyrValAlaAspLysValPheAsnPheAlaGlyGluArg 181
      |||
      |||
      |||
      |||
505 GCGAAGAAACCGCGCTTACCAAGCATACAGAGGCTCAAAACATCAT 554
      |||
      |||
      |||
      |||
182 ..MetGlyAsnLeuValProSerHisIgaLysrSerPheAlaLeuAl 197
      |||
      |||
      |||
      |||
555 CAAGCCCTGGCTGGGGGAAACACCATGCTCCGCGGACCAAGCTTC 604
      |||
      |||
      |||
      |||
197 aArgGluLeuLeuLysGlyGlyGlyValGlyValLeuValAsp..... 211
      |||
      |||
      |||
      |||
605 CCTCCCTCAAGAAGCGGAGAGCGGTATGGGTGATTTCTTCGGCAAA 654
      |||
      |||
      |||
      |||
212 .....GlnLysPheGlyLysGlyLeuThrThrLysPhePheGlyLeu 225
      |||
      |||
      |||
      |||
655 CCTGCTTACCATGACCGCTGCGGCAAAATTGGACACAGCTCAAAAGCGT 704
      |||
      |||
      |||
      |||
226 GluValAlaGThrAsnProLeuAlaLysLeuValAlaG..... 238
      |||
      |||
      |||
      |||
705 GAAACCCCTGTTTCTGC.....TGCAGACGCTGC 736
      |||
      |||
      |||
      |||
239 .....GlnPheAsnCysAspValTyrProAlaIargLysIleArgLeuP 253
      |||
      |||
      |||
      |||
737 CTGGCGGACAAGTTTGCATATTCACATCCGCGCTTCGACAGCAATGTG 786
      |||
      |||
      |||
      |||
253 roAspAsnArg...TyrArgLeuGluIleGluPro...LysValGluLe 267
      |||
      |||
      |||
      |||
787 .....AACGGCAGCAAAAGCCCATGATGCGCGCGCTGT 818
      |||
      |||
      |||
      |||
268 ProAlaGAspArgInLysGlyAsnValAspIleGlnAlaThrAlaGlnLeu 284
      |||
      |||
      |||
      |||
819 CAACCGCAATGCCGATATTTGGATACGCCGTTTCCGACGACGATCTGT 868
      |||
      |||
      |||
      |||
284 uAsnAspLysValGlnSertTrrValaIargGluTyrProGlyGlnTrrPleu 301
      |||
      |||
      |||
      |||
869 TTATGTACACCGCTAC 885
      |||
      |||
      |||
      |||
301 rPtyrHisAspArgTyrP 306
      |||
      |||
      |||
      |||
seq_name: p1r2:F82350
seq_documentation_block:
lipid A biosynthesis (kds)2-(lauroyl)-lipid IVA acyltransferase VC0212 [imported] - Vibrio
Species: Vibrio cholerae

```

```

C:/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:/Accession: F82350
R:/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
1, R.R.; Makalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:/Reference number: A82035; MUID:20406833
A:/Accession: F82350
A:/Status: preliminary
A:/Molecule type: DNA
A:/Residues: 1-325 <HEI>
A:/Cross-references: GB:AE004111; GB:AE003852; NID:99654614; PIDN:AAF93388.1; GSPDB:GN
A:/Experimental source: serogroup O1; strain N16961; biotype El Tor
C:/Genetics:
A:/Gene: VC0212
A:/Map position: 1

alignment_scores:
      Quality: 141.00      Length: 186
      Ratio: 1.270      Gaps: 6
Percent Similarity: 59.677      Percent Identity: 25.269

alignment_block:
US-09-303-518D-569 x F82350 ..

Align seg 1/1 to: F82350 from: 1 to: 325

319 GCGTGGGACATGTGCAGCAGGCGTTTGACAAACAGGAAGGCGTCTATT 368
    ||| |||:::|||||:::|||||:::|||||:::|||||:::
122 G1GUG1UG1uasnleuProleuLeuAspserG1G1uArgValIlele 138
369 CATTACAGCCGCACATGCGCAGCTACGATTGGCGGAGCGCTACATAGGCC 418
    :::::|||||:::|||||:::|||||:::|||||:::
138 ulcValProInserTrrPAlleAsprTgAlaAlaValMetLeuAla 155
419 AGCAGCTTCGCTCCGCTGCAGCCGCATGTACAAACCGCGGAATCAAA 468
    :::::|||||:::|||||:::|||||:::|||||:::
155 laArg..G1YTyrLySValAlaAsnIleMetLySProGInlArgAsnPro 170
469 GCGATACACAAATCATGCAGCGCGCGCGGTTGCGGCGCAAGGAAAC 518
    ||| :::::|||||:::|||||:::|||||:::
171 lAlaAsprTrpleuMetHisValGlnArgMetGlnTyrG1yAlaTyr 187
519 CGCGCCCTACAGCATACAGGGGTCAAACAAATCATCAAGCCCTCGGT 568
    ||| :::::|||||:::|||||:::|||||:::
187 epHe...ThrArgLusergLyIleLySProPheLeuArgSerIleGln 203
569 CGGCGGAGCAACCATCGTCTGCCCGCAGCAGTCCCTCCCTCAAGAA 618
    |||| :::::|||||:::|||||:::|||||:::
203 ergLyTyrValG1YTyrTrrPAlProAsprG1uAsnHisG1yProGIn 218
619 GCGCGGGAAGGCGTATGGTGCGATTCTTTCGCGCAACCTGCGTATCCAT 668
    :::::|||||:::|||||:::|||||:::|||||:::
219 .....AsnserValrPheValProPheAlaIhrG1uLySAlaIhrPle 233
669 GAGCGTGGCGCAAAATTCGACACGTC...AAAGCGTGAAAAACCTGT 715
    :::::|||||:::|||||:::|||||:::|||||:::
233 ulyGg1rPhe1yLySMeAlaLySLeuCySlyAlaHisValValrProL 250
716 TTTTCTGTGTGCGAACGCCGTGCGCTGGCGGACAAAGGTTTCGATTTCACATC 765
    :::::|||||:::|||||:::|||||:::|||||:::
250 eumetserCySlyrAsnserAsprserG1yArg...TyrG1uValHisIle 265
766 CGCCCC..GTCCAAAGGGGAATTGAACGGCGACAAAGCCCATGATGGCGC 812
    ||| :::::|||||:::|||||:::|||||:::|||||:::
266 leuProAlaLeuGlnAsnPrPheProThrG1yAspr1uG1uAlaPaAla 282
813 CGGTGTCAACCGCAATGCCGAAATATGGAATCGCGCGTTTTCAGACGAGAT 862
    ||| :::::|||||:::|||||:::|||||:::|||||:::
282 uAlMetserAsnArgAlaIleG1uAlaLeuValThrProGInProG1uGInT 299

```

863 ATCTGTTT 870  
 299 yrmctrr 301  
 seq\_name: p1r2:c82350

seq\_documentation\_block:

lipid A biosynthesis lauroyl acyltransferase WC0213 [Imported] - Vibrio cholerae (strain C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C:Accession: G82350  
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardson, D.; Esmolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, H. 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A:Reference number: AB2035; MUID:20406833  
 A:Accession: G82350  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-318 <HER>  
 A:Cross-references: GB:AE004111; GB:AE003852; NID:99654614; PIDN:AAF93389.1; GSPDB:GN001  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 A:Gene: WC0213  
 A:Map position: 1

alignment\_scores:

Quality:	136.50	Length:	317
Ratio:	0.858	Gaps:	15
Percent Similarity:	50.158	Percent Identity:	24.921

alignment\_block:

US-09-303-518D-569 x G82350 ..

Align seg 1/1 to: G82350 from: 1 to: 318

```

58 TTGACCGCCCTGCTCAATGCTCCTGCTGCGCTTCCTGCTGCA 107
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
32 LeuLeuAlaLeuLeu.....ValasnLeuProTyrProValLeuLe 46
108 CACGCTGGGAAACCGCTCGGA..... 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
46 ulsrlleglYarlglyLeuGlyGlnPheSerMetArpHeqlyLysIys 63
130 .....CATCTGGCGCTTTTACCTTTTAAG 153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63 rGValHisIleAlaArgArGAsnLeuGluLeuAlaPheProThrMetSer 79
154 GAAGACCC.....GGCGCATGTCGCCCAATATGCGTCAGGACGCAT 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80 GlnSerGluIleAspAlaPheValLeuGlnAsnPhelyAsnThrGly.. 95
198 GAATCCGACCCCAAAACGGTCAAGCGCTTTTTCGGAAGGCAAAAG 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
96 .....AlaAlaLeuIleGluThrGlyIleT 104
248 GCGGTTTGAACTGCCCCCGCTTTTTCAGAAAACCGGAACATAGAA 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
104 hTTPPrPheTrrProThrTrrPheLysArGlyIleLeuIleAspLysasp 120
298 ACAATGTTCAACGCGTACACGCGTGGAAACATGTGACAGAGCTTGA 347
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 Thr.....GlnAlaIleAr 125
348 CAACAC.....GAGGCGCTATTCATCAGCCGCACATCG 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
125 gGlnHisAlaLysThrGlyGlnGlyValLeuLysCysValHisAla 142
386 GCAGCTACGATTGGCGGACGCTACATCAGCCAGCAGCTTCGCTCCG 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
142 euAsnLeuGluIleThrAlaArgAlaPheAla...ValLeuGlyIleGly 157

```

seq\_name: p1r2:T43023

seq\_documentation\_block:

ceoa protein - Burkholderia cepacia  
 C:Species: Burkholderia cepacia  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 11-Jan-2000  
 R:Charon, R.L.; Barry, J.J.; Burns, J.L.  
 submitted to the EMBL Data Library, April 1997  
 A:Reference number: Z22290  
 A:Accession: T43023  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-405 <CHAF>  
 A:Cross-references: EMBL:U97042; NID:92109269; PID:92109270; PIDN:AAB58160.1  
 A:Experimental source: strain K61-3  
 A:Gene: ceoa

alignment\_scores:

Quality:	133.50	Length:	212
Ratio: <td>1.362 <td>Gaps:</td> <td>10</td> </td>	1.362 <td>Gaps:</td> <td>10</td>	Gaps:	10
Percent Similarity: <td>46.226 <td>Percent Identity: <td>28.774</td> </td></td>	46.226 <td>Percent Identity: <td>28.774</td> </td>	Percent Identity: <td>28.774</td>	28.774

alignment\_block:

US-09-303-518D-569 x T43023 ..

```

436 CTGACCGCATGTACAAACCGCCGAATAATCAAGCATAGCAAAATCAT 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
158 GlyTyrGlyValTyrArgProHisSerAsnProAlaTyrGluPheIleG 174
486 GCAGCGGGGAGGTTCCGGCAAGAAACCGCGCTACACCATAC 535
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
174 nTyrArgGlyArgThrArgAsnGlyAsnGlnLeuIleAsnArgThr... 189
536 AAGCGTCAACAAATCATCAAGCGCTGTCGGGCGGAACCAATCAT 585
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
190 ..AspIleLysIleMetIleArgValLeuArgGlnGlyValArgLeuPhe 205
586 CTCTCGCCCGAACCATGCTCCCTCCCTCAGAAAGCGGG.....GAAG 629
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 TyrLeuProAsp.....GlnAspTyrGlyHisAsnLysSe 217
630 GGTATGGGTGATTTCTTCGCG...AAACCTGCTATACATGACGCTGC 676
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
217 rValPheValProPhePheAlaValGluGluAlaCysThrThrThrGlyT 234
677 CGGCAAAATTTGACACAGTCACAAAGCGTGAAACCGCTGTTTCTGCTGC 726
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
234 hrSerIleLeuAlaTyrThrSerHisCysAlaIleValIle..... 247
727 GAACGCTGCTGCGCGCAAGATTTCGATTTGCACATCCGCGCGTCA 776
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
248 .....GlySerGlyPhe.....ArgAsnAlaG 255
777 AGGG.....GAATTGAACGCGCACAAAGCCCATGAT..... 807
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
255 nGlyArgTyrGluIleMetAlaAspLysSerIleGluAlaAspTyrProg 272
808 .....GCCGCGTGTCAACCGCAATGCGCAATAT 837
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
272 InLysAspGluThrAlaAlaAlaAlaTyrMetAsnLysPheValGluGlu 288
838 TGGATGACCGCGTTTTCGAGCGAGTATCTGTTATGACAAACCGCTACAA 867
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
289 IleIleLeuArgAlaPheGluGlnTrrPheMetLysIleLysArgPhe 305
888 A 888
305 s 305

```

Align seg 1/1 to: T43023 from: 1 to: 405

```

150 AAAGGAGACCGCGCGATCGTCGCCAATATGCGTCAAGCAGCATGA 199
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
161 LysArgGluProGluGlyArgArgSerArgAlaGlyAsnGlyAlaHisG 177
200 ATCCGACCCCAAAACGGTCAAGCCGTTTGGCGAAGCGCAAAAGCC 249
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
177 nSerArgLeuTyr.....AlaHisHisA 185
250 GCTTGGACATCGCCCGCGGTTTTCAGAAAACCGAGACATAGAAC 299
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
185 rGaLaGlyValGlyProArgValAlaArg.....GlyAsnHisAlaArg 199
300 AATGTTCAACCGGTACACGGCTGGAAACATGTGA.....GCAGGCTT 343
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
200 GlnArgArgValGlyArgArgValGlyArgAlaAlaAspAspAlaGly 216
344 TGA.....CAACACGAAAGGCT.....G 363
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
216 eGlyValAlaAspLeuArgValValArgArgArgAlaAspLeuProA 233
364 CTAATCATCAC..... 374
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
223 IalleHisGlnArgArgAlaGlnArgProGlnGlyAlaGlyArgAlaArg 249
375 .....GCCGACATCGGACGCTACGATTTGGGCG 403
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
250 ProArgGluArgAsnArgLeuLeuAlaGlnArgArgArgPheGlyArg 266
404 GACGCTACATGACCGACGCTCCGTCGCCGCGCATGTACAAA 453
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
266 G.....GlnProAlaArgHisValVal.....ArgHisAspProA 278
454 CGCGCAAAATCAAGCGATAGCAAAATCATGCAGC..... 491
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
278 rGaLaArgProLeuArgGlnArgGlyArgHisProGlyProGlyProLeu 294
492 .....GGGCAAGGTTCCGGCAAGGAAAGAAACCGCGCTACGACATAC 535
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
295 ArgThrArgGluGlyArgGlnArgGlnArgAlaAlaArgGlyAlaAlaArg 311
536 AAGGGGTCAACAATCATCAAGCCCTGCGTCCGGCGAAGCAACATC 585
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
311 gArgArgGlyArgGlnHisArgProGlyGlnGlyValArgValAlaArgArg 328
586 GTCTGTGCCCGACCGCTCCCTC..... 608
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
328 rGProAlaGlyProArgValAlaSerArgSerAlaAlaArgAspAlaAla 344
609 .....CCTCAAGAAAGCGCGGGAAGCGGTATGGGT 638
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
345 ArgGlnProAlaArgAspArgGlnArgAlaValGly 356

```

seq\_name: p1r2:T05722

```

seq_documentation_block:
  licheninase (EC 3.2.1.73) - barley
  N:Alternate names: Licheninase
  C:Species: Hordeum vulgare (barley)
  C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 08-Oct-1999
  C:Accession: T05722
  R:Slakecki, N.; Baulcombe, D.C.; Devos, K.M.; Ahluwalia, B.; Doan, D.N.P.; Fincher, G.B.
  M:J. Gen. Genet. 224, 437-449, 1990
  A:Title: Structure and tissue-specific regulation of genes encoding barley (1->3, 1->4)
  A:Reference number: S13734; MUID:91094782
  A:Accession: T05722
  A:Status: Preliminary; translated from GB/EMBL/DBJ
  A:Molecule type: DNA
  A:Residues: 1-335 <SLAB>
  A:Cross-references: EMBL:X56260; NID:g18991; PIDN:CAA39700.1; PID:g18992
  A:Experimental source: cv. NK 1558
  C:Genetics:

```

A:Map position: 1L  
A:Introns: 25/2  
C:Keywords: glycosidase; hydrolase

alignment\_scores:  
Quality: 133.00 Length: 359  
Ratio: 0.842 Gaps: 18  
Percent Similarity: 44.011 Percent Identity: 23.955

alignment\_block:

US-09-303-518d-569 x T05722 ..

Align seg 1/1 to: T05722 from: 1 to: 335

```

81 CTCGCGCGCGCGCTTCGTCGACACGCGTGGAAACCGGCTCGAC 130
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
10 LeuAlaLeuAlaLeuLeuGlyAlaPheAlaSerIleProGlnArg.. 25
131 ATCTGCGCTTTTACCTTTTAAAGAAACCGCGCGCATCGTCCCAAT 180
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
26 .ArgGlyVal.....HisArgGlyValLeuArgHisGlnArgGlnGln 40
181 ATGCGTCAGGACGACAT..... 197
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
40 eAlaGlyGlyGlnHisArgGlyGlnHisValGlnValGlnArgAspGln 56
198 .....GAATCCCGACCC.....CA 211
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
57 LeuHisAlaAlaValAlaArgSerArgProGlyGlyAlaAlaGlyArg 73
212 AAACGGTCAAAAGCGCTTTTGGGAAACCGCAAAAGCGGCTTT..... 254
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
73 gHisGlyArgGlnArgCysCysGlyArgAlaGlnArgArgAlaLeuGlnP 90
254 ..... 254
90 rArgArgGlnSerArgSerGlyCysIleValGlyGlnGlnGlnHisPro 106
254 ..... 254
107 GlyValProGlnGlyLeuLeuProValCysLeuArgGlyGlnArgGlyArg 123
255 .....GGAATTCGCGCGCGCTTTTTCAGAAACCGGGA 288
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
123 gArgArgArgHisProGlnProCysProArgHisGlnGlnArgAlaGly 140
289 GA.....CATAGAAACATGTTCAAGCGCTACAGCGCTGGGAGACA 329
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
140 rGaLaGlyLeuArgArgAlaGlyProHisGlnGlnArgHisValGly... 155
330 TGTGCACAGCGCTTTGCAACACGAAAGCGCTGATTCATCAGCCGCG 379
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
156 ...ValAlaGlnHisProGlyGlyValGlnProAlaValAlaArgValLe 171
380 ACATCGGACGACGATTTGGGCGGACCGCTACACGCGCTGGGACGA 429
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
171 HisHisArgArg.....GlyGlyArgValHisGlyProArgGlyAlaVal 185
430 TTCGCGTGCACCGGCTGATCAAAACCGCGCAAAATCAAGCGATAGCA 479
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
185 AlProCysPro...HisArgArgAlaAlaHisGlnGlnHisLeuProVal 200
480 AATCATCGAGGCGGCGGCGGTTCGCGCAAAAGGAAACCGCGCGCTAGCA 529
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
201 ProGlyLeuGlyLeuGlnProGlnArgHis..... 210
530 GCATCAAGGGGTCAACAATCATCAAGCCCTGCGTGGGCGAAGCA 579
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
211 .....GlyHisGlyLeuArgGlyAlaLeuHisArgLeuAla 221
580 ACATATGTT...CTGCGCGCA.....CCAGCTCCGCTC..... 608
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

```

221 rghIsArgGlyProGlyArgLeuLeuArgValProGluProValArgHis 237
609 ..... CCTCAAGAGCGGGGAA 628
238 HisArgGlyArgLeuLeuHisGlyHisGlyGlnAlaArgArgLeuGlnArg 254
629 GCGTATGGGTGATTTCTTCGCAACCTGCCTATACCATGACGCTGGG 678
254 gGlnAlaGlyGlyValArgGlnArgValAlaValSerArgArgHisGlyG 271
679 GCAAAATTGGCACAGCTCAA..... 698
271 LysArgProGlnArgGlnAspLeuGlnProValProHisGlnProArg 287
699 AGCGCTGAACACCTGTTTTCGTGCGAGACGCTGCCTGGCGGCAAG 748
288 ArgAlaArgHisPro.....Pro.ProProGlyArgH 298
749 GTTTCGATTTGCACATCCGCCGTCACAGGGGAATTGAACGCGACAA 798
298 IsArgAspLeuArgLeuLeuHisValGlnArgGlnProGlnGlyGlnArg 314
799 ..... GCCCATGATGCCGCTGTTCACACCCCATGCCGAATTATGAT 842
315 ArgGlyAlaGlnLeuGlyAlaLeuLeuProGlnHisAla..... 327
843 ACGCCGTTTCCGACGACGATATCTG 867
328 AlaArgLeuProHisGlnLeuLeu 335

```

seq\_name: p1r2:E90949

```

seq_documentation_block:
  heat shock protein Mabb [imported] - Escherichia coli (strain O157:H7, substrain R1MD 05
  C/Species: Escherichia coli
  C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
  C/Accession: E90949
  R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
  gasawara, N.; Yasunaga, T.; Kunara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
  DNA Res. 8, 11-22, 2001
  A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
  A/Reference number: A9629; MUID:21156231; PMID:11258796
  A/Accession: E90949
  A/Status: preliminary
  A/Molecule type: DNA
  A/Residues: 1-323 <NAV>
  A/Cross-references: GB:BA000007; PIDN:BA835988.1; PID:q13362033; GSPDB:GN00154
  A/Experimental source: strain O157:H7, substrain R1MD 0509952
  C/Genetics:
  A/Gene: ECs2565

```

alignment\_scores:

```

  Quality: 129.50      Length: 297
  Ratio: 0.881        Gaps: 11
  Percent Similarity: 49.495  Percent Identity: 23.22

```

alignment\_block:

US-09-303-518d-569 x E90949 ..

Align seg 1/1 to: E90949 from: 1 to: 323

```

64 GCCCGCTCAATAGCTCTCCCTGCTGCCGCTTTCCTGTGCACAGCT 113
   |||  ::::: |||  |||  ::
31 AlaAlaMetAlaGlyIleAlaLeuThrProProLysPheArgAspProI 47
114 GGGAACCGGCTCGACATCTGGCGTTTACCTTTAAAGGAAGACGGCG 163
   :  |||||  ::::: |||  |||  |||  ::::: |||
47 eleuAlaArgLeuGlnArgIleAlaGlyArgLeuGlyLysSerSerArg 64
164 CGCGCATGTGCGCAATATG..... 183
64 rGArgAlaLeuIleAsnLeuSerLeuGlyPheProGluArgSerGluAla 80

```

```

184 ...CSTCAGCGAGCGCATGATCCGACCCCAAAAGCGTCAAGCGTTT 230
   |||  ::|||
81 GluArgGluAla.....IleValAspGluMetPh 90
231 TGGGAAACCGCAAAAGCGTTTGGAACTTCCCGCGCTTTTTCAGAA 280
   |||  ::::: |||  ::|||
90 eAlaThrIleProGlnAlaMetAlaMetMetAlaGluLeuAlaIleArg 107
281 AACCGAAGACATAGAACATGTTCAAAGCGGTACAGCGTGGACAT 330
   |||  |||  ::|||
107 LyrProGluLysIleGlnProArgValAsp...TrpGlnGlyLeuGlnIle 122
331 GTGCAGCAGGCTTTGGCAACACGAAAGCGCTGCTATTCACACGCGCA 380
   ::::: |||  ::|||
123 IleGlnGluMetArgArgAsnAsnGlnLysValIlePheLeuValProH 139
381 CATCGGCGAGTACGATTTGGCGGACGCTACATCAGCAGCGTTCCGT 430
   |  ::::: |||  ::|||
139 SGlyTrpAlaValAspIleProAlaMetLeuMetAlaSerGln...Gly 155
431 TCCGCTGACCCCATGTACAAACCGCGAAATCAAGGATGACAA 480
   ::::: |||  ::|||
155 LInLysMetAlaIleMetPheHisAsnGlnGlnAsnProValPheAspTyr 171
481 ATCATGACAGCGGCGAGGTTCCGCGCAAGAAACCGCGCTACAG 530
   |||  ::|||
172 ValTrpAsnThrValArgArgArgPheGlyGlyArgLeuHisAlaArg 188
531 CATCAAGGGGTCAACATCATCAAGCGCTGCTGGCGGCAAGCA 580
   :  ::|||
188 n...AspGlyIleLysProPheIleGlnSerValArgGlnGlyTyrTrp 204
581 CCATGCTCTGCC.....GACCACTGCCCTCCCTCAAGAGCGGG 624
   |||||  |||||  |||||
204 LyrTyrTyrLeuProAspGlnAspHisGlyProGlnHisSerGln..... 218
625 GAAGCGCTATGCGGTGATTTCTTCGCAACCTGCCTATCCATGACGCT 674
219 .....PheValAspPhePheAlaThrTyrLysAlaThrProAl 232
675 GCGCGCAAAATTGGCACACGCTCAAGCGCTGAACCCCTGTTTTCGCT 724
   ::|||  |||
232 AlIleGlyArgLeuMetLysVal.....CysA 241
725 GCGAAGCCCTG.....CTGGCGGACAGCT..... 750
241 rGAlaArgValIleProLeuPheProIleTyrAspGlyLysThrHisArg 257
751 TTCGATTTGCACATCCGCGCTCAAGGGGA...TTGACGGCGGCA 797
   ::  ::::: |||||  ::|||
258 LeuThrIleGlnValArgProPheMetAspAspLeuGlnAlaAspAs 274
798 AGCCCATGATCCGCGGTGTTCACACGCAATGCGGATATTTGATACGCG 847
   ||  ::|||  ::|||
274 PHisThrIleAlaArgArgMetAsnGlnGluValIlePheValGlyP 291
848 GTTTCGAGCAGATCTGTATGTACACCGCTACAA 888
291 roArgProGlnGlnTyrThrTrpIleLeuLysLeuLeuLys 304

```

seq\_name: p1r2:A85798

seq\_documentation\_block:

```

  suppressor of htrb, heat shock protein [imported] - Escherichia coli (strain O157:H7,
  C/Species: Escherichia coli
  C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
  C/Accession: A85798
  R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
  Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallanite, E.; Potamousts, K.; Apoda
  Nature 409, 529-533, 2001
  A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
  A/Reference number: A85480; MUID:21074935; PMID:11206551

```

A:Accession: A85798  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-323 <STO>  
 A:Cross-references: GB:AE005174; NID:g12515911; PIDN:AG56845.1; GSPDB:GN00145; UWGP:229  
 A:Experimental source: strain 0157:H7, substrain EDL933  
 C:Genetics  
 A:Gene: msdb

alignment\_scores:  
 Quality: 129.50 Length: 297  
 Ratio: 0.981 Gaps: 11  
 Percent Similarity: 49.495 Percent Identity: 23.232

alignment\_block:  
 US-09-303-518D-569 x A85798 ..  
 Align seg 1/1 to: A85798 from: 1 to: 323

```

64 GCGCTGCTCAATGCGCTGCTGCGCTTTCTGTCACAGCT 113
||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
31 AlaAlaMetAlaGlyIleAlaLeuThrProGlyPheArgSerPro11 47
114 GGGAAACGGCGCTGACATCTGGCGTTTACCTTTAAGAAAGACGGCG 163
: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
47 eLeuAlaArgLeuGlyArgIleAlaGlyArgLeuGlyLysSerSerArg 64
164 GCGGATCGTCGCGCCATATG..... 183
||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
64 rGAlaAlaLeuIleAsnLeuSerLeuGlyPheProGlyArgSerGluAla 80
81 GluAlaArgGluAla.....IleValaAspGluMetPh 90
231 TCGCGAAACGGCAAGCGGTTTGAAGCTGGCCCGCTTTTCAGAA 280
||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
90 eAlaThrAlaProGlnAlaMetAlaMetMetAlaGluLeuAlaIleArg 107
281 AACCGGAGACATAGAAACAATGTTCAAGCGGTACAGCGCTGGAGACAT 330
||| ||| ||| ::| ::| ::| ::| ::| ::| |||
107 LyrProGlnLysIleGlnProArgValaAsp...TrpGlnGlyLeuGlnIle 122
331 GTGCAGACAGCTTTGGACAAACAGAGCGCTGATTTTCATCAGCCGCA 380
::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
123 IlegIleGlnMetArgArgAsnAsnGlnLysValIlePheLeuValProH1 139
381 CATCGGAGCTACGATTTGGCGGCGGCTACATCAGCCAGCGCTTCGCT 430
| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
139 sGlyTrrAlaValaAspIleProAlaMetLeuMetAlaSerGln...GlyG 155
581 CCATGCTGCTGCGC.....GACACAGTCCCTCCCTCCAGAGAGCGGG 624
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
204 LyrTrrTrrLeuProArgGlnAspHisGlyProGlnHisSerGlu..... 218
625 GAAGCGGTATGGGTGATTTTTCGCGCAACAGCTGATACACAGCT 674
::| ||| ||| ||| ||| ||| ||| ||| ||| |||
219 .....PheValaAspPhePheAlaThrTrrTrrLysAlaThrLeuProAl 232
675 GCGGCAAAATTTGGACAGCTCAAGAGCGTGAAACCCCTGTTTCTTGCT 724

```

```

232 AlieGlyArgLeuMetLysVal.....CysA 241
::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
725 GCGAAGCGCTG.....CCTGGCGGACAGGT..... 750
241 rGAlaArgValAlaProLeuPheProIleTrrAspGlyLysThrHisArg 257
::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
751 TTCGATTTGCATCCGCCCGCTCCAGAGGGAA...TTGAACGGCGACAA 797
258 LeuThrIleGlnValAlaArgProPheMetAspAspLeuGluAlaAspAs 274
798 AGCCCATGATGCCCGCGTGTTCACACCGCAATGCCAATATTGGATACGC 847
||| ::| ::| ::| ::| ::| ::| ::| ::| ::|
274 rHisThrIleAlaArgArgMetAsnGlnGluValGluIlePheValGlyP 291
848 GTTTTCCGACGACATCTGTTATGTATACACCGCTACAAA 888
||| ||| ||| ||| ||| ||| ||| |||
291 rArgProGlnGlnTrrThrTrpIleLeuLysLeuLeuLys 304

```

seq\_name: p1r2:j00405

seq\_documentation\_block:  
 hypothetical 119.5k protein (uvrA region) - Micrococcus luteus  
 N:Alternate names: ORF 1 protein  
 C:Species: Micrococcus luteus, Micrococcus lysodeikticus  
 C:date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 04-Feb-2000  
 C:Accession: J00405  
 R:Shiota, S.; Nakayama, H.  
 Mol. Gen. Genet. 217, 332-340, 1989  
 A:title: Micrococcus luteus homolog of the Escherichia coli uvrA gene: identification  
 A:reference number: S04781; MUID:89364717  
 A:Accession: J00405  
 A:Molecule type: DNA  
 A:Residues: 1-1106 <SH1>  
 A:Cross-references: EMBL:X15867  
 A:Note: this reading frame extends between two stop codons and does not begin with a  
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo

alignment\_scores:  
 Quality: 129.00 Length: 373  
 Ratio: 0.915 Gaps: 19  
 Percent Similarity: 37.802 Percent Identity: 23.324

alignment\_block:  
 US-09-303-518D-569 x J00405 ..

Align seg 1/1 to: J00405 from: 1 to: 1106

```

21 GCGTTTCCCGCTTTGGCAACCGCATGTCACATCTGTGACCGCGCTGC 70
||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
202 AlaAlaLeuProAlaValaArgArgAlaGlyLysProAlaAspProAlaH1 218
71 TCANAATG.....CCTGCTCCCTGCGG 90
||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
218 AspArgGlyProAlaArgGlyAlaAlaArgAlaHisProLeuProGlyA 235
91 CCGGTTTCCGTGCT..... 104
||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
235 IalArgAlaArgGlyProArgProGlnGlyArgValaArgGlyProValaGln 251
105 .....GCACAGCTGGGAAACCGGCTCGACATCTGGCGCTTTACCTTT 148
::| ||| ||| ||| ||| ||| ||| ||| ||| |||
252 GlyProValaHisAlaGlyLeuArgArgArgGlyArgGlyLysArgProAl 268
149 TAAAGAGACCGCGCGCG.....CATGTCGCG 177
||| ||| ||| ||| ||| ||| ||| ||| |||
268 AleuGlyProAlaArgAlaGlnGlnAlaGlnAlaHisHisArgArgA 295
178 AATATGCTGACGACGACATGA..... 200
::| ::| ::| ||| ||| ||| ||| ||| |||
285 rGArgGlyProArgHisGlnGlyLysHisProProAlaProHisArg 301

```

```

201 .....TCCCGACCCCAACGGTCACG 223
302 LeuGlyGlyAspArgProGluAlaGlyGlyArgProArgArgGlyArgGly 318
224 CCGTTTTCGGGAAAGCGCAAAAGCGGTTTGGAAGCTCCCGCGCGTTT 273
318 LArgGlyArgGlyAlaGlyArgArg..... 326
274 TTCGAAACCGGAAACATAGACATGTCGAAAGCGGTACAGCGG 323
327 ..GluGlyGlnGlnGlnHisArgGlyValArgArgAlaGlyArgArg 341
324 GGAACATGTGCAGACGAGCTTTGGACAA.....ACACG 355
342 .....GlyGlnProProValProLeuValLeuArg 351
356 AAGGCGTGTATTCATCAGCGGCAATCGGACGATTCGATTTGGCGGA 405
351 GGUAlaLeuLeuProGlnArgAlaArgAlaAspArgGlyArgAspArg 368
406 CG.....CTACATCAGCCAGACGCTTCGTTCC...GCTGAC 440
368 LAlaLeuValLeuLeuGlnGlnProValArgArgValProArgValHis 384
441 CGGCATGTTCAAACCGCCGAAATATCAAGCGCTAGACAAATCATGCAAG 490
385 ArgHis.....ArgLeuProProAlaGly 392
491 CGGAGCGGTTTCGCGCAAGAAACCGGCTTACAGCATACAGG 540
392 yGlyProGlyProArgArgArgGlnArgAlaValPro...AlaArgG 408
541 CTCGAA..... 545
408 LArgArgArgAlaValAlaAlaArgGlnValHisLeuGlyLeuLeuAla 424
546 .....ACAAATCAT... 554
425 AlaArgAlaArgArgAlaGlyGlnGlyAspArgLeuLeuProGlyHisPr 441
555 ....CAAGCCCTCGGTCGCGGCAAGCAACCATGCTGCGCGGA.... 596
441 oValGlnGlyProAspArgGlyGlyAlaArgArgProAlaArgGlnG 458
597 .....CCAGTCCCTCCCTCAAGAG... 620
458 LLeuGlnGlyGlyAspValArgProValArgProArgAlaPro 474
621 .....CGGGAAGCGGTATGCGGATTCCTGCGGCAAC 655
475 LeuHisHisGlyLeuArgGlyArgHisPro.....LeuArgAspAl 488
656 CTCGCTATACATGACGCTGCGGCAAAATTCGACACAGCTCAAGCGGTG 705
488 a.....GlnAlaArgGlyAspArgValGlyArgArgProArg.... 500
706 AAAACCGTGTTCGCTGCGAAGCGCTGCTGCGGACAGTTTCGA 755
501 .....AlaLeuArg 503
756 TTTCGACAT.....CCGCCCG 772
504 ValValHisAlaGlyAspProValProGlyValProArgGlyProProGly 520
773 TTCAGAGGGAATTCACCG 791
520 nProHisGlyProGlnArg 526
seq_name: p12:A42608
seg_documentation_block:
(Kdo)2-(lauroyl)-lipid IV4 acyltransferase (EC 2.3.1.-) - Escherichia coli

```

```

N:Alternate names: membrane-bound lytic transglycosylase
C:Species: Escherichia coli
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A42608; I54979; G64947
R:Karow, M.; Georgopoulos, C.
J. Bacteriol. 174, 702-710, 1992
A>Title: Isolation and characterization of the Escherichia coli msbB gene, a multicop
A:Reference number: A42608; M01D:92121107
A:Accession: A42608
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-323 <KAR>
A:Cross-references: GB:M87660; NID:9146876; PIDN:AAA96706.1; PID:9146877
A:Experimental source: strain K-12, substrain W3110
R:Engel, H.; Smink, A.J.; van Wijngaarden, L.; Kock, W.
J. Bacteriol. 174, 6394-6403, 1992
A>Title: Murein-metabolizing enzymes from Escherichia coli: existence of a second lyt
A:Reference number: I54979; M01D:93015688
A:Accession: I54979
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-323 <RES>
A:Cross-references: GB:M87660; NID:9146876; PIDN:AAA96706.1; PID:9146877
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; M01D:97426617
A:Accession: G64947
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-323 <BLAT>
A:Cross-references: GB:A8000279; GB:U00096; NID:91788134; PIDN:AAC74925.1; PID:917881
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: msbB; mlt
C:Function:
A:Description: transfers myristate or laurate, activated on ACP, to (Kdo)2-(lauroyl)-
A:Pathway: lipopolysaccharide core biosynthesis
A>Note: decanoyl, palmitoyl, palmitoleyl and (R)-3-hydroxymyristoyl-ACP are poor acy
C:Keywords: acyltransferase; lipopolysaccharide core biosynthesis; transmembrane prot
F:23-43/Domain: transmembrane #status predicted <TM>
F:85-105/Domain: transmembrane #status predicted <TM>
F:133-153/Domain: transmembrane #status predicted <TM>

```

alignment\_scores:

	Quality:	127.50	Length:	297
	Ratio:	0.867	Gaps:	11
	Percent Similarity:	49.495	Percent Identity:	23.232

alignment\_block:

```

US-09-303-518D-569 x A42608 ..
Align seg 1/1 to: A42608 from: 1 to: 323

```

```

64 GCCCTCTCAATGCTCTCCCTGCGGCTTCCTGCTGCACACGCT 113
||| :::::||||| ||| :::::
31 AlaAlaMetAlaGlyLeuAlaLeuThrProProLysPheArgAspPro11 47
114 GGAACACGCGCTCGACATCTGCGCTTTACCTTTAAAGAACCGCG 163
: |||||:::||||| ||| |||||:::
47 eLeuAlaArgLeuGlyArgPheAlaGlyArgLeuGlyLysSerSerArg 64
164 CGCGCATCGTCGCAATATG..... 183
64 rArgArgAlaLeuIleAsnLeuSerLeuCysPheProGlnArgSerGlnAla 80
184 ...CGTCAGCAGAGCATGATCCGACCCCAAAACGTCAAAGCGGTTT 230
||||:||||| ||| :::::
81 GlnArgGlnAla.....IleValAspGlnMetPh 90

```







Align seg 1/1 to: S07132 from: 1 to: 886

```

51 CATCTGTTGACCCG..... 65
21 HSPROTHrGLyArGThrGLyLeuGLyAspArgLeuLeuIleAlaIle 37
66 .....CCTGTCMAATGCTCTCCCTGCGCGCTTCTGTCGACACG 111
37 tValThrIaGLInAlaProIleAlaIleAlaValPheAsnGLnProSerGly 54
112 CTGGCAAAACGGCTCGACATCTGGCTTTTACCTTTAAAGAGACCG 161
54 IAlaIaArgAlaPheGLnProMetThrAlaIaProThrGLnGLyAla 70
162 CGC.....GGCATTGTCG 175
71 ArgIleAlaProAlaIleGLnLysGLnAsnLeuLeuAlaArgAla 87
176 CCAATATGCGTCAGCAGCATGAAATCCGACCCCAAAACGGTCAAAGCC 225
87 nGLyPheLeuAspGLyGLyAspGLnArgAlaGLnProAlaGLnAla 104
226 GTTTTGGGGAACGGCAAAAGCGCTTGGAACTTGGCCCGCTTTT 275
104 IYArgAlaGLYHISGLYArgArgGLYArgArg..... 115
276 CAGAAACCGGAGACATGAAACAT..... 302
116 ArgArgGLyGLYArgArgArgGLyGLYArgPheLeuLeuGLyLe 132
303 .....GTTCAAAGCGGTACACGGCTGGAAACATGTGCAGCAG 339
132 vAlaValArgProAlaGLYArgGLYGLnThrLeuGLY.....ProAlaG 147
340 GCTTTGGCAAAACAGA..... 356
147 LysrGLYLysIleArgGLYGLnProPheLeuArgArgIleAlaIle 163
357 .....AGGCTGCTATTCATCAGCCGCG 379
164 ILeGLnGLYAspLeuGLYHISpHeArgSerAlaValAlaThrGLY 180
380 ACATCG.....CAGC 390
180 nHISGLnMetProIleAlaIleAlaLeuAlaGLYAlaLeuAspGLY 197
391 TACGATTTGGCGAGCAGTACAT...CAGCCGACGTTCCCTCCCT 437
197 rGLArgArgGLYArgGLYHISGLYGLYAlaIleAlaGLYAla 213
438 GACCGC...CATGTACAAACCGCGAAATCAAGC..... 470
214 AspHISGLYHISValAlaGLYAlaIleArgHISAlaLeuLeuPheGL 230
471 .....GATAGACAAATCATGACGGCGGCGAG 498
230 uGLYAlaValAlaPhePheIleAspAspAspGLnAlaGLnIleGLY 247
499 GTTCCGGCAAGAAACCGCGCTACAGTACATACAGGGTCAAACA 548
247 LysGLnLysIleArgArgSerArgProHISHISArgArg..... 260
549 AATCATCAAAAGCCTGCTGGCGGAGCAACATGCTGTCGCCGACG 598
261 .....ProArgGLnGLY.....HISArgProProSerPr 270
599 AGCTCC.....CTCCCTCAAGAAAGCGGCGA. 626
270 oProProArgArgArgAlaHISLeuArgMetProLeuHISArgSerGLY 287
627 .....AGCGTATGAGTGA..... 641
287 IaGLnAlaValaGLYAlaIleProArgProLeuGLYGLnGLYAspPhe 303

```

```

641 ..... 641
304 GLYGLnGLnHISGLnSerLeuProProAlaIaGLnGLYAspArg 320
642 .....TTTCTTCGCG 651
320 gLeuGLnIleAspLeuGLYLeuAlaArgAlaGLYAspProPheGLnGL 337
652 AATCGTCCATACATGACGCGCGCAAA..... 683
337 LInThr.....GLYAspSerLeuValYsArgArg 346
684 .....ATTGGACACGTCAAGGCGTGA 706
347 AlaGLnAlaValaArgArgArgLeuLeuIleGLY...ArgGLnGLnArg 362
707 AAACCTGTTTTCG.....CTGGCAACGCTGCTGCGGACAAAGT 750
362 rGLYProValaGLYIleGLYLeuAlaArgProPheIleArgThrGLY 379
751 TTCGATTTCACATCCG.....CCCGCTCA 776
379 IsArgLeuGLnGLnProGLYLeuGLYHISArgArgHISHisArgArgPro 395
777 AGGGGA..... 782
396 HISGLYArgPheProArgGLnIleGLYAlaGLYGLnGLYGLnThrGLY 412
783 .....ATTGAAGGGGCAAAACCATGATGCGCGCTGTCACAGCGCA 827
412 tValValYsAsnArgGLnAsnProLeuAlaIaGLYAlaGLnProArg 428
seq_name: p1r2:C75580

```

```

seq documentation block:
adenine deaminase-related protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
R:White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: AF5250; M01D:20036896
A:Accession: C75580
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <WHI>
A:Cross-references: GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF12376.1; PID:g646
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0268
A:Map position: 2

```

```

alignment_scores:
Quality: 116.00 Length: 268
Ratio: 0.906 Gaps: 13
Percent Similarity: 47.761 Percent Identity: 26.866

```

alignment\_block:

US-09-303-518D-569 x C75580 ..

Align seg 1/1 to: C75580 from: 1 to: 376

```

60 GACCGCCCT...GCTCAATGCTCTCCCTGCGCGCTTCTCGTGC 106
|||||
120 AspArgProThrPvalSerArgArgLeuArgAlaIaIaArgLeuAlaIn 136
107 ACACGCTGGAAACGGCGTACATGCGCTTTTAAAGAGAA 156
:
:
:

```

```

136 r.....LeuArgglyAlaGlyAsnLeuArgglyArg 148
157 GACCGCGCGCATGTCGCCAATATCGCAGGAGGCGATGATCCGA 206
148 LysErlAlaIleArgArgAlaProProAlaArgTyrArgArg 164
207 CCCCMAAGGTCMAAGCGCTTTTGGGMAAGCGCGGTTTG 256
165 ProArgAlaIleArgLeuGlyArgGlyHisLeuArgSerAlaArgAlaLeuAl 181
257 AACTGGCCCCGGCTTTTCAGAAACCGAACAATAGAACATGTTTC 306
181 ahIAlaIleProAlaIle...SerArgProAlaArgHisArg..... 193
307 AAGCGGTACACGCGTGGGAACATGTGCAGCAGCTTTGGACAAACGA 356
194 .....AlaCysGlyAlaGlyHisArg 201
357 AGGCTGCTATTCATCAGCCGACATCGGAGTACGATTTGGCGGAC 406
202 ArgProAla.....GlyArgArgArgProLeuArgAlaArgArg 215
407 GTCACAT...CAGCCAGACGCTCCGCTCCGCTGACCCGCTTACAAA 453
215 lValGlyLeuLeuAspValGlyHisArgProAlaArgArgHis..... 229
454 CCGCCGMAATCAAGCGATAGACAATCATCAGCGCGCGCGGTTGC 503
230 ..ProGlyHisGlyHisSerAlaArgArgAlaSerGlyGlyProAlaArg 245
504 CGCAGAGMAAACCAGCGCTACACGCTACAGGCTCAACAAATCA 553
246 ArgGlyHisArgGlyProAla.....ArgGlyGlySerGlyAl 258
554 TCAGACCGCTGGCTGGCGGAGCAACATCGCTCCGCGCA..... 596
258 aArgThr.....AlaArgArgArgHisArgProGlyArgGlyArg 272
597 .....CCAGTCCCTC.....CCG 611
272 rGgIyProArgAlaValAlaProAlaLeuArgGlySerAspArgPro 288
612 TCAGAGAGCGGCGGAGCGCTATGGGATTTCTTGGCAACCTGGCT 661
289 lAlaSerGlyArgGlyArgArgCysProGly..... 298
662 ATACCATGACGCTGGCGCAAAATTCGACACAGTCMAAGCGGTAAAAAC 711
299 .ProGlyAspGlyGlyAlaLeuAlaGly..... 308
712 CTGTTTTCGCTGCGAAGCGCTGCTGGCGGACAAAGTTTCGATTGCA 761
309 .....LeuHisLeuAlaLeuSerArgHisHisProGlyHisSer 321
762 CATCGCGCGCGTCAGAGGGAATTGAACGGCGACAAAGCCCATGATGCG 811
332 ArgProGlyArgAspSerGlyAlaGlyAlaAspAlaIleArgProAlaIle 338
812 CCGT 815
338 gArg 339
seq_name: p1r2:E86965
seq_documentation_block:
conserved_hypothetical_protein_MU0453 [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: E86965
R:Cooley, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001

```

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;  
A:Title: Massive gene decay in the leprosy bacillus.  
A:Reference number: A86909; MUID:21128732; PMID:11234002  
A:Accession: E86965  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-320 <STO>  
A:Cross-references: GB:AL450380; NID:g13092692; PIDN:CAC29961.1; GSPDB:GN00147  
A:Gene: MU0453

alignment\_scores:  
Quality: 115.50 Length: 253  
Ratio: 0.808 Caps: 14  
Percent Similarity: 56.522 Percent Identity: 28.063

alignment\_block:  
US-09-303-518D-569 x E86965 ..

Align seg 1/1 to: E86965 from: 1 to: 320

```

75 ATGCTCTCCCTGCTGCGCTTC...CTGTGCAACGCTGGGAACC 121
||||| ||||| ||| :..... |||
37 MetProGlySerAlaIleArgPheAlaPheAsnThrGlyAlaIleGlyAl 53
122 GGCTGGACATCTGGCGTTTACCTTTAAAGAACCGCGCGCATC 171
||||| :..... |||
53 aAlaArgAsnGlyGlyProAspGlyHisLeuArgGlyAlaIle 69
172 GTGCCCAATATGCGTCAGCAGCATGAATCCGACCCCAAAAGGTCA 221
70 lIleGly.....ValGlyAlaIleAspValProAsp...AlaLeuIleAr 83
222 AGCGGTTTTCGGAAGCGCAAAAGCGGTTTGGAACTTCCCGCGGT 271
... :..... |||
83 gThrSerLeuAlaSerGlyAlaArgTyrTrpArgGlyAla..... 96
272 TTTTCAAGAAACCGGA...GACATGAAACAATGTTCAAGCGTACAC 318
||||| ||| :..... |||
97 ..PheArgLeuProThrMetAsnProHisThrValAlaIleGlyAla 112
319 GGCTGG.....GAACATGTGCAGCAGGCTTGGCAAAACGCA 356
... :..... |||
113 AlaTyGlyIleAspLeuHisArgValGlyAlaLeuAspAlaGlyAla 129
357 AGCGTCTATTCATCAGCGCGCACATCGGACGCTACGATTTGGCGGAC 406
||||| :..... |||
129 gGlyValIleLeuAlaLeuProHisSerGlyAsnThrAspMetAlaGly 146
407 GTCATCAGCGAG...CAGCTTCCGTTCCCGCTGACCGCC...ATGTAC 450
... :..... |||
146 alTrpLeuValGlyHisIleGlyThrPheThrValAlaIleArgLeu 162
451 AAACCGCGGAATTC.....AAAGCATAGACAAATATCATCGACGGG 494
||||| :..... |||
163 LysProGlySerLeuTyArgArgPheIleAsnTyArgGlyHisLeuGly 179
495 C...AGGTTTCGGCGCAAGGAAAAACCGCGCTACGACGATCAAGGG 541
| :..... |||
179 yPheGlyValIleuProMetSerAlaAsnLysAspThrGlyPheGlyAla 196
542 TCAGAACAAATCATC...AAAGCGTGGCTGGCGGAGCAACATCATGTC 588
... :..... |||
196 eupHegIleuLeuGlyGlyAlaGlyHisArgSerHisArgValIleCysLeu 212
589 CTGCCGACACGCTCCCTCCCTCAAGAAAGCGGGAAGCGGTATGGT 638
... :..... |||
213 MetAlaIleArg.....AspLeuThrArgThrGlyValGlyAla 225
639 GGATTTCTTGGCAACCTGCTATACATGACGCGGCGGCAAAATTCG 688
||||| :..... |||
225 lAspPhePheGlyGlyProThrArgMetProAlaGlyProAlaLysLeu 242

```

